Electronic Journal of Probability

Vol. 8 (2003) Paper no. 4, pages 1–93.

Journal URL
http://www.math.washington.edu/~ejpecp/
Paper URL
http://www.math.washington.edu/~ejpecp/EjpVol8/paper4.abs.html

STATE DEPENDENT MULTITYPE SPATIAL BRANCHING PROCESSES AND THEIR LONGTIME BEHAVIOR

Donald A. Dawson [1]
School of Mathematics and Statistics, Carleton University
Ottawa, Canada K1S 5B6
ddawson@math.carleton.ca
http://lrsp.carleton.ca/directors/dawson

Andreas Greven [2]
Mathematisches Institut, Universität Erlangen-Nürnberg
Bismarckstraße D-91054 Erlangen, Germany
greven@mi.uni-erlangen.de
http://www.mi.uni-erlangen.de/ greven

Abstract

The paper focuses on spatial multitype branching systems with spatial components (colonies) indexed by a countable group, for example $\mathbb{Z}^d$ or the hierarchical group. As type space we allow continua and describe populations in one colony as measures on the type space. The spatial components of the system interact via migration. Instead of the classical independence assumption on the evolution of different families of the branching population, we introduce interaction between the families through a state dependent branching rate of individuals and in addition state dependent mean offspring of individuals. However for most results we consider the critical case in this work. The systems considered arise as diffusion limits of critical multiple type branching random walks on a countable group with interaction between individual families induced by a branching rate and offspring mean for a single particle, which depends on the total population at the site at which the particle in question is located.

The main purpose of this paper is to construct the measure valued diffusions in question, characterize them via well-posed martingale problems and finally determine their longtime behavior, which includes some new features. Furthermore we determine the dynamics of two functionals of the system, namely the process of total masses at the sites and the relative weights of the different types in the colonies as system of interacting diffusions respectively time-inhomogeneous Fleming-Viot processes. This requires a detailed analysis of path properties of the total mass processes.

[1] Supported by an NSERC grant and a Max Planck Award.
[2] Supported by a DFG-Schwerpunkt

In addition to the above mentioned systems of interacting measure valued processes we construct the corresponding historical processes via well-posed martingale problems. Historical processes include information on the family structure, that is, the varying degrees of relationship between individuals.

Ergodic theorems are proved in the critical case for both the process and the historical process as well as the corresponding total mass and relative weights functionals. The longtime behavior differs qualitatively in the cases in which the symmetrized motion is recurrent respectively transient. We see local extinction in one case and honest equilibria in the other.

This whole program requires the development of some new techniques, which should be of interest in a wider context. Such tools are dual processes in randomly fluctuating medium with singularities and coupling for systems with multi-dimensional components.

The results above are the basis for the analysis of the large space-time scale behavior of such branching systems with interaction carried out in a forthcoming paper. In particular we study there the universality properties of the longtime behavior and of the family (or genealogical) structure, when viewed on large space and time scales.

Keywords and phrases: Spatial branching processes with interaction, multitype branching processes with type-interaction, historical process, universality, coupling of multidimensional processes, coalescing random walks in singular random environment.

AMS subject classification (2000): Primary 60K35; Secondary 60G57, 60J80

Submitted to EJP on March 11, 2001. Final version accepted on January 30, 2003.

Contents

Introduction: Background and program

Branching systems with a *spatial* structure and involving migration have been studied extensively in the literature and we refer here to Dawson 1993 [D] for a list of references ranging from Branching Random Walks, Interacting Feller Diffusions to Superprocesses. The analysis of all branching systems depends crucially on the independence of the evolution of the subpopulations descending from one particular individual. We want to remove this *independence assumption* and to instead replace it with a state dependent branching mechanism satisfying exchangeability properties. As a consequence we shall find some new phenomena in the longtime behavior.

More specifically we would like to understand the large time and large space-scale behavior of spatial branching systems with migration with two additional features, namely *with interaction between families* and of *multitype* populations. The treatment of both these points require some new techniques and we view the model described as a necessary first step. The second step involves dropping the assumption of criticality of the branching and introduction of a logistic drift term. The overall objective is to find in sufficiently large classes of evolutions the *universal* pattern of the longtime behavior. The longtime behavior in focus is that of the spatial system itself as well as that of the *historical process*, which describes the statistics of the genealogies and their spatial structure.

Various interactions between families in a branching evolution have been considered in the literature, examples are so-called multilevel branching systems (see Greven and Hochberg (2000) for a recent survey and references), mutually catalytic populations (see Dawson and Perkins (1998) [DP]) and competing super processes (Evans and Perkins (1998) [EP]). The interaction between families of branching populations we have in mind in this paper, is that the branching rate of an individual *depends* on the *total population* at its location and that also the mean offspring may depend on the total population. There are the extreme cases of unlimited resources leading to a branching rate independent of the population at the site (classical branching random walk) and the other case of finite resources leading to a branching rate, which vanishes beyond a certain level of the population per site, an example are systems of the type of the Moran model (Fleming-Viot in the diffusion limit) with fixed total population. The first three main tasks (handled in this work in Theorems 1 - 4 and Propositions 0.1 - 0.5) are to obtain in the multitype situation with interaction between families a

- unique characterization and construction of the process as well as the corresponding process describing the genealogy of individuals called historical process (Theorems 1, 2)
- description of the longtime behavior of the process and the historical process (Theorem 3, 4)
- description of the dynamics of total mass and relative weights of the types and to give the clan decompositions of the infinitely old population (Propositions 0.1 - 0.5).

We shall briefly describe some of the questions and outline the content before we describe for motivation the particle model leading to our diffusion model. We model the state of a population at a site by a measure on the generic type space $[0, 1]$, which means that the components are multi and typically even infinite-dimensional so that the problem of existence and unique description of the model arises and is of a serious nature (recall that for resampling systems (Fleming-Viot) the problem of defining the process uniquely is unsolved for general state dependent rates). We construct the above mentioned model, which are systems with *countably many interacting* (via migration) *measure valued diffusions*, via *martingale-problems*. We show in particular that these martingale problems are well-posed. This is all formulated in section 0(a). Furthermore we describe in 0(c) the evolution of the corresponding total mass and relative weight processes. In addition we define via a martingale problem in section 0(d) the so-called historical process, which records effectively for a fixed time t the evolution through times s before t of the individual "mass" alive at time t and that of its relatives.

In 0(b) we discuss the longtime behavior of the process and in 0(d) the longtime behavior of the historical process. The processes constructed above exhibit a dichotomy in behavior depending on properties of the symmetrized migration similar to that of classical spatial branching processes. Moreover, further effects for the relative weights of types may occur depending on the branching rates for small and large populations.

In spite of the multidimensional (and interdependent) nature of the components, it is possible to construct *equilibrium states* in the case of transient symmetrized migration and to prove theorems about the *domain of attraction* of these equilibria. For this purpose we first introduce the process of the total masses in the components and second the process of the relative weights of the various types. These objects can be studied using new types of *couplings* and *dual processes*, the latter are processes in a space-time random environment with singularities. The information on the equilibrium state is completed by means of the limiting state (of an infinitely old population) for the historical process associated with the equilibrium process of the system.

In the case of the recurrent symmetrized migration, we find *clustering* of the total masses (i.e. small islands with large masses and most of the space empty) but the composition of types raises some new questions. Namely, are there certain cases in which in the long term the local picture is *monotype* while in other cases it is *multitype*? Furthermore we can ask whether the rare spots with large populations have a monotype or a multitype structure. We show that these questions lead to the study of interacting diffusions in randomly fluctuating media, which by duality reduces to coalescing random walks in a space-time-inhomogeneous medium. This leads to new very challenging problems.

The next goal (for a forthcoming paper), is to show and explain that many features of the *long time* and *large space scale behavior* are *universal* and do not depend on the particular form of the branching rate and are even present in models with resampling instead of branching (Fleming-Viot process). This is true first of all for the qualitative behavior of the law of the system as time tends to infinity. However, we have in mind for this forthcoming paper an analysis of the universality of *finer properties* of the longtime behavior. For this end we have in this current work in chapters 2 - 4, derived some facts needed later on as a technical tool.

The analysis of the questions indicated above is complicated and does not always give explicit results. The analysis will therefore be based on a scheme of *renormalization*, the so-called *multiple space-time scale analysis*.

0. The model and results on longtime behavior

Motivation: A particle model The measure-valued processes we consider should be viewed as diffusion limits of particle systems. Since this will be needed as a technical tool and since the model itself and questions on its longtime behavior are easier to grasp we start by describing this particle model.

Think of particles located on a countable group I which represents the *colonies* in which the population is organized (for example the hierarchical group Ω_N or $\mathbb{Z}^d$) and in addition these particles have a specific type. In addition to that structure we can introduce the notion of a *family* (sometimes also called *clan* in the literature) consisting of all the descendents of one particular individual present at time 0.

This way the state of the system can by described by

$$\eta(u,\xi) = \#\text{ particles of type } u \text{ in } \xi, \qquad u \in [0,1],\ \xi \in I.$$

In other words we can describe the configuration by a product (over I) of atomic measures on [0,1], with weights in $\mathbb{N}$.

We obtain "maximal“ multitype systems by assigning to every family a type as follows. Let the particles present at time 0 be randomly labelled as follows: with each individual we associate as a label (called type) one random variable taken from an i.i.d. collection of [0,1]-valued uniformly distributed random variables. Each of the particles being present initially will give rise to a distinct family as time evolves. Every multitype model can be encoded in the model where every particle initially has a different type.

This construction allows us also to observe spatial extension and density of the family founded by an initial particle during the evolution.

Denote with $\widetilde{\mathcal{M}}_a \subseteq \mathcal{M}$ the set of atomic measures with weights in $\mathbb{N}$ and $\mathcal{M}$, the set of measures, is endowed with the weak topology of measures. Then write using above notation our process in the form:

$$X(t) = (x_\xi(t))_{\xi \in I}, \quad x_\xi \in \widetilde{\mathcal{M}}_a([0,1]). \tag{0.1}$$

The evolution on the space of measures is now induced by the particle model evolving as follows:

(i) Every particle moves independently of every other particle according to a continuous time random walk with transition kernel $\bar{a}(\xi,\eta)$ on $I \times I$, (here $\bar{a}(\xi,\eta) := a(\eta,\xi)$, the latter we use later as the basic parameter).

(ii) Every particle splits into two particles or disappears at rate (recall $x_\xi(t) \in \widetilde{\mathcal{M}}_a([0,1])$)

$$h[x_\xi(t)([0,1])]. \tag{0.2}$$

In the critical case births and deaths occur with equal probabilities (or rates). Otherwise a birth in colony ξ of each type u occurs with the biased probability:

$$\frac{1}{2} + c_\xi(u, x_\xi(t)([0,1])). \tag{0.3}$$

The different particles (and families) in a given colony interact through their contributions to the total mass of the colony. However conditioned on the complete total population process, all particles act independently.

(iii) The transitions in (i) and (ii) are independent.

Very important functionals of this process are the *total number of particles* in a colony and the *relative weights of types* in a colony. It is possible to describe the evolution of these functionals in a nice way for the total masses, but for the relative weights conditioned on the total mass process such a nice description is only possible after taking the diffusion limit.

We can associate a *historical process* with such a multitype branching system with interaction between families. We fix a time t and consider the total population of particles alive at that time. To every particle we associate the path tracing back to his father, then continue the path by tracing further back to the grandfather etc. This way we can form the locally finite measure on paths of length t by giving every path of descent starting from a particle alive at time t a unit weight. By extending the path as a constant path before time 0 and after time t we obtain measures on $D((-\infty,\infty),\Omega)$. If we now vary t we obtain a stochastic process and this process is called the historical process.

In the branching particle model just described we can pass to the continuum mass limit. Namely, by giving every particle mass ε, increasing the number of initial particles at rate ε^{-1}, speeding up the branching rate of a particle by the factor ε^{-1} and scaling h by $h(\varepsilon x)$, we obtain as $\varepsilon \to 0$ the *diffusion limit*, which is the object we discuss later on in this paper (in the case of $c(\cdot,\cdot)$ and $h(\cdot)$ constant this process is usually called super random walk). One of the first problems in chapter 0(a) (Theorem 1) is to define the limiting object uniquely by a martingale problem. Since our populations are multitype, hence multidimensional this is actually nonstandard. Finally, construction and analysis of the *historical process* in the diffusion limit is in chapter 0(d) (Theorem 3).

(a) The model and two functionals of the basic process. In this subsection we first construct and uniquely characterize in Theorem 1 our process (having countably many interacting measure valued components) and then define in a second paragraph some characteristic functionals.

(i) Construction of the process via a martingale problem.

In order to define the model of a system of interacting Feller's continuous state branching systems with additional family-interaction through a state dependent branching rate, we need to introduce now first the ingredients.

(α) We use as labels for the spatial components of the system a general countable group and we abbreviate it with Ω. The group action is denoted $+$, the inverse $-$ and we use the term translation invariant and ergodic for invariance and ergodicity with respect to the group action. By ergodic average over a random field we mean the average over an increasing sequence of finite subsets such that the ergodic theorem is valid (see [K] for the necessary conditions). Two choices for Ω are of particular interest, namely $\mathbb{Z}^d$ and the hierarchical group Ω_N.

(β) To define the state space of the process we first need the set of *types* which is $[0,1]$ with the metric topology (only the separability is important). Denote by $\mathcal{M}([0,1])$ the set of (nonnegative) finite measures on $([0,1], \mathcal{B}([0,1]))$.

We shall define *the state space* $\mathcal{E}$ of the process, which is a subset of

$$(\mathcal{M}([0,1]))^{\Omega}, \tag{0.4}$$

by the following device, which prevents explosion in this spatially *infinite* system in finite time. Let γ be a strictly positive function on Ω, which is summable and which we specify more concretely in (δ) below. Then define for $X = (x_\xi)_{\xi\in\Omega},\ x_\xi \in \mathcal{M}([0,1])$:

$$\|X\|_\gamma = \sum_\xi \gamma(\xi) x_\xi([0,1]). \tag{0.5}$$

The set $\mathcal{E} = \mathcal{E}_\gamma$ is given by

$$\mathcal{E}_\gamma = \{X \in (\mathcal{M}([0,1]))^{\Omega} | \quad \|X\|_\gamma < \infty\}. \tag{0.6}$$

We use on $\mathcal{E}$ the *topology* given by the product topology, where for each component we have the topology of weak convergence of measures.

(γ) Next we describe the *migration mechanism* between colonies. The results of chapter 0 hold if we use any irreducible random walk on Ω with transition rates denoted $a(\xi,\eta) = a(0,\eta-\xi)$ and satisfying $\sum_\xi a(0,\xi) < \infty$. Since we have assumed that all sites communicate, given two points ξ,η there is a sequence $\xi_1,\dots,\xi_k$ of minimal length,

$$\rho(\xi,\eta) \quad \text{with} \ \ a(\xi,\xi_1) > 0, \dots, a(\xi_k,\eta) > 0. \tag{0.7}$$

The function ρ defines a metric on Ω.

(δ) We shall choose in the sequel a function γ in the definition of $\mathcal{E}$ in (β) above which matches $a(\cdot,\cdot)$, namely choose as *state space* $\mathcal{E} = \mathcal{E}_\gamma$, with γ satisfying:

$$\sum_{\xi\in\Omega} \gamma(\xi) a(\xi,\xi') \le M\gamma(\xi') \quad \text{for some} \quad M \in \mathbb{R}^+. \tag{0.8}$$

In particular this implies that $\sum_{\xi'\in\Omega} a(\xi,\xi') x_{\xi'} < \infty$ for $x \in \mathcal{E}$. The function γ can be constructed as follows.

Denote the k step transition-probability of the random walk by $a^{(k)}$. Take any function β on I such that $\beta > 0$ and $\sum\limits_\xi \beta(\xi) = 1$ and some $M > 1$ sufficiently large so that the following series converges:

$$\gamma(\xi') = \sum_k \sum_\xi M^{-k} \beta(\xi) a^{(k)}(\xi,\xi'). \tag{0.9}$$

(ε) We shall need a function $h : [0,\infty) \to \mathbb{R}^+$ the future state dependent *branching rate* satisfying

$$\begin{aligned} &(i) \quad h(x) > 0 \quad \forall x \in (0,a) \text{ with } a \in \bar{\mathbb{R}}^+ \\ &\qquad\quad h(x) = 0 \quad \forall x \ge a \\ &(ii) \qquad h \text{ is locally Lipschitz on } [0,\infty) \\ &(iii) \qquad\quad h(x)/x \underset{|x|\to\infty}{\longrightarrow} 0. \end{aligned} \tag{0.10}$$

From this function h a continuous function g (which will be the *diffusion function* for the total mass process of a component) is derived by putting

$$g(x) = h(x)x \quad \text{for} \quad x \in [0,\infty). \tag{0.11}$$

(ζ) Finally we need a term describing the *deviation from criticality* of type u at total population size x:

$$c : \Omega \times [0,1] \times [0,\infty) \to \mathbb{R}, \quad c \text{ is measurable} \quad \text{and} \quad \sup_{\xi,u,x} |c_\xi(u,x)| < \infty. \tag{0.12}$$

The classical super random walk is obtained by setting the functions c and h constant. The classical Fleming-Viot process is then also embedded in this model by taking $a = 1,\ c = 0$ and $h(x) = x(1-x)$.

The process we shall consider will take values in $\mathcal{E}$ and will be defined by a *martingale problem*. In order to define the corresponding generator, consider two classes of smooth functions F on $\mathcal{E} \subseteq (\mathcal{M}([0,1]))^{\Omega}$ given by:

$$F(X) = \prod_{i=1}^{n} g_i\big(\langle x_{\xi_i}, f_i\rangle\big), \qquad \xi_i \in \Omega,\ f_i \in C([0,1]),\ g_i \in C_b^2(\mathbb{R}),\ n \in \mathbb{N}, \tag{0.13}$$

$$F(X) = \int f(u_1,\dots,u_n) x_{\xi_1}(du_1)\dots x_{\xi_n}(du_n), \text{with} f \in C^{+}([0,1]^n),\ n \in \mathbb{N} \tag{0.14}$$

In both cases, $\mathcal{A}$, the generated algebra is an algebra of smooth functions on $\mathcal{E}$. In addition in the first case, (0.13), the functions are smooth and bounded giving a *measure determining algebra*. In the second case the algebra is measure-determining provided that the corresponding moment problem is well-posed. In particular, the second class will be used later when we are dealing with the probability measure-valued case in which the moment problem is automatically satisfied.

On $\mathcal{A}$ we define differential operators. Functions $F \in \mathcal{A}$ are viewed as restrictions of functions on products of signed measures and hence derivatives in direction δ_u are defined by:

$$\frac{\partial F}{\partial x_\xi}(X)(u) = \lim_{\varepsilon \to 0} \frac{F(X^{\varepsilon,\xi,u}) - F(X)}{\varepsilon}, \tag{0.15}$$

where ($\delta_{\eta,\xi}$ is the Kronecker symbol with $\delta_{\xi,\eta} = 0$ or 1 depending on whether $\xi \neq \eta$ or $\xi = \eta$ and δ_u denotes the unit mass on $u \in [0,1]$):

$$X^{\varepsilon,\xi,u} = (x_\eta + \varepsilon \delta_{\eta,\xi}\delta_u)_{\eta \in \Omega}\,. \tag{0.16}$$

The second derivative $\frac{\partial F}{\partial x_\xi \partial x_\eta}$ in direction δ_u, δ_v is defined as $\frac{\partial}{\partial x_\xi}\left(\frac{\partial F}{\partial x_\xi}(X)(u)\right)(v)$.

Since $\frac{\partial F}{\partial x_\xi}$ is weakly continuous for $F \in \mathcal{A}$ and since finite measures can be approximated in the weak topology by linear combinations of δ-measures it suffices for first order derivatives to consider the directions δ_u, respectively δ_u, δ_v for second order derivatives.

Finally we define the following *differential operator* G acting on functions $F \in \mathcal{A}$ (see (0.13)):

$$\begin{aligned}(GF)(X) \quad = \quad & \sum_{\xi}\Bigg[\int_0^1 \frac{\partial F}{\partial x_\xi}(X)(u)\Big\{\sum_{\xi'} a(\xi,\xi')(x_{\xi'}(du) - x_\xi(du))\Big\} \\ & + \int_0^1 \frac{\partial F}{\partial x_\xi}(X)(u) h(x_\xi([0,1])) c_\xi(u, x_\xi([0,1])) x_\xi(du) \\ & + \int_0^1\int_0^1 \frac{\partial^2 F}{\partial x_\xi \partial x_\xi}(X)(u,v) h(x_\xi([0,1])) x_\xi(du)\delta_u(dv)\Bigg], \quad F \in \mathcal{A}.\end{aligned} \tag{0.17}$$

Before we come to the main result of this subsection recall the notion of the solution of a martingale problem.

Definition 0.1. Define for every random variable X with values $\mathcal{E}$ the solution of the (G,X)-*martingale problem* w.r.t. $\mathcal{A}$ recall (0.13), as a law P on $C([0,\infty),\mathcal{E})$ such that almost surely $X(0) = X$, the random variable X is independent of the process $(X(t) - X(0))_{t \geq 0}$ and

$$\left(F(X(t)) - F(X(0)) - \int_0^t (GF)(X(s))ds\right)_{t \geq 0} \tag{0.18}$$

is a P local martingale w.r.t. the canonical filtration for all $F \in \mathcal{A}$. $\square$

At this point we can give a rigorous definition of the diffusion limit of the particle model described in (0.1) using the key result:

Theorem 1. *The* (G, X_0)*-martingale problem w.r.t. to* $\mathcal{A}$ *is well-posed for every* $X \in \mathcal{E}$. $\square$

Remark The local martingale in (0.18) can be replaced by a martingale, if h is a bounded function.

Definition 0.2. The $\mathcal{L}((X(t))_{t \geq 0})$ is the solution of the (G, X_0)-Martingale problem and is referred to as *interacting system of branching diffusions* with *branching rate h* and *bias c*. $\square$

Using a result of [DGV] Proposition 1.2, [D] Theorem 8.2.1 and our Proposition 0.4 yields that the states of the process are the atomic measures denoted $\mathcal{M}_a$:

Lemma 0.1. *For* $t > 0$ *the states are atomic measures a.s., one even has*

$$P(x_\xi(t) \in \mathcal{M}_a([0,1]), \quad \forall \xi \in \Omega, \ t > 0) = 1. \quad \square \tag{0.19}$$

(ii) Total mass and relative weights processes

The system just defined can be analyzed by looking at the process of total mass in a colony and then after conditioning on that process for all colonies it remains to study the process of the relative weights of the single types.

We will use the following notation for the *total mass process*:

$$\bar{X}(t) = (\bar{x}_\xi(t))_{\xi \in \Omega}, \quad \bar{x}_\xi(t) = x_\xi(t)([0,1]). \tag{0.20}$$

More generally we will abbreviate with $\bar{\mu}$ for every $\mu \in \mathcal{M}([0,1], \mathcal{B})$, the quantity $\mu([0,1])$.

We shall see that for the initial conditions, which we study later on, whenever $X(0) \not\equiv \underline{0}$,

$$P(\bar{x}_\xi(t) = 0) = 0, \tag{0.21}$$

but in general, i.e. depending on $h(0)$ and $X(0)$, $P(\bar{x}_\xi(t) > 0, \ \forall t \in \mathbb{R}^+)$ is not 1. However in our context, it is always the case that the set of 0's has Lebesgue measure 0 as we shall see below in subsection 0(c).

Once the total masses are given the important quantity is the *relative weight* of the different types in the population at site ξ. Define the following stochastic process, called *relative weights process*, by using the convention $0/0 = 0$, uniquely as:

$$\widehat{X}(t) = (\hat{x}_\xi(t))_{\xi \in \Omega} = ((x_\xi(t)/(x_\xi(t)([0,1])))_{\xi \in \Omega}. \tag{0.22}$$

Later in subsection 0(c) we give simple descriptions of the dynamics of these functionals, which are straightforward for total masses but need some preparation and a new class of processes to describe the relative weights process *conditioned* on the complete total mass process (see Proposition 0.1, 0.4, (0.57), (0.58)).

(b) The longtime behavior. In the rest of the paper we assume that our branching mechanism is *critical* that is

$$c.(\cdot) \equiv 0. \tag{0.23}$$

The longterm behavior exhibits the competition of the two parts of the dynamics. The migration tends to push the system towards constant states and the diffusion part (i.e. branching part) tends to push it toward local extinction and dominance of one type. The question is here as in other interacting systems (branching random walk, voter model, etc.) who wins.

In order to understand this feature of the longtime behavior of our system we first return to the system of total mass processes (0.20). This will turn out in 0(c) to be a system of interacting diffusions. The behavior, as time tends to infinity, of this system of interacting diffusions is well-known ([S, 92], [CG1, 94], [CFG, 96]) and depends primarily on the transition kernel and not on the function h. As usual put $\widehat{a}(\xi,\eta) = \frac{1}{2}(a(\xi,\eta)+a(\eta,\xi))$. Then the longtime behavior of the total mass process depends on whether the diffusive part dominates the behavior or whether the migration parts dominates, which translates into whether the symmetrized kernel $\widehat{a}(\cdot,\cdot)$ is recurrent or transient. In the first case we get local extinction, in the second we get convergence to nontrivial equilibria. A variant of this is true for the process $X(t)$ in the transient case while in the recurrent case a new type of behavior is exhibited ((d) below), where in fact the behavior of h near 0 and ∞ plays an important role.

We can identify the extremal invariant measures ((a) below), identify the behavior of $\mathcal{L}(X(t))$ as $t \to \infty$ ((b) below) and we can say something about finer properties of the longtime behavior ((c),(d) below). We say a translation invariant measure has finite intensity if the ergodic average of the field $\{x_\xi([0,1]), \xi \in \Omega\}$ is a.s. finite. In this case we can define with a general version of the ergodic theorem an intensity measure, usually denoted with the letter θ.

Theorem 2. *Assume here that $h(x) > 0$ for all $x > 0$.*

(a) *Case: $\widehat{a}(\cdot,\cdot)$ is transient.*
There exists for every finite measure θ on $[0,1]$ a measure ν_θ on $(\mathcal{M}([0,1]))^\Omega$ with:

$$\begin{aligned} &(i) && \nu_\theta \text{ is translation invariant and ergodic} \\ &(ii) && \nu_\theta \text{ is an extremal invariant measure of the evolution} \\ &(iii) && E_{\nu_\theta} x_\xi = \theta \\ &(iv) && \nu_\theta \text{ is concentrated on } (\mathcal{M}_a([0,1]))^\Omega, \text{ where the } a \text{ denotes atomic.} \end{aligned} \tag{0.24}$$

The measure ν_θ is uniquely determined by (i) and (iii). The set of all those translation invariant measures with finite intensity, which are invariant for the evolution, is the closed convex hull of the set

$$\{\nu_\theta,\ \theta \in \mathcal{M}([0,1])\}. \tag{0.25}$$

Case: $\widehat{a}(\cdot,\cdot)$ is recurrent.
The only invariant measure which is translation invariant and has finite intensity is $\delta_{\underline{0}}$, where $\underline{0}$ denotes the product over Ω of the zero-measures on $[0,1]$.

(b) *If the process starts in a translation invariant and ergodic initial law with finite intensity measure θ then:*

$$\mathcal{L}(X(t)) \underset{t\to\infty}{\Longrightarrow} \begin{matrix} \nu_\theta & \text{if} & \widehat{a}(\cdot,\cdot) \text{ is transient} \\ \delta_{\underline{0}} & \text{if} & \widehat{a}(\cdot,\cdot) \text{ is recurrent,} \end{matrix} \tag{0.26}$$

in the weak topology of probability laws, based on the product of the weak topology of $\mathcal{M}([0,1])$, with the metric topology used on $[0,1]$.

(c) *If $\widehat{a}$ is transient and if $\mathcal{L}(X(0))$ is translation invariant and ergodic, then if we define $\bar{\nu}_\theta$ and $\widehat{\nu}_\theta$ as the law of $\bar{X}$ respectively $\widehat{X}$ under ν_θ (recall(0.20), (0.22)) we have:*

$$\mathcal{L}(\bar{X}(t)) \underset{t\to\infty}{\Longrightarrow} \bar{\nu}_\theta, \quad \mathcal{L}(\widehat{X}(t)) \underset{t\to\infty}{\Longrightarrow} \widehat{\nu}_\theta. \tag{0.27}$$

The measure $\nu_{\bar{\theta}}$ does only depend on $\bar{\theta}$, has mean $\bar{\theta}$ while $\widehat{\nu}_\theta$ has mean $\widehat{\theta}$ but does depend on both $\bar{\theta}$ and $\widehat{\theta}$.

(d) *If $\widehat{a}$ is recurrent then we distinguish two cases. Assume initial states which have a finite intensity. Condition on the complete total mass process. It is possible to construct for every fixed t a time-inhomogeneous random walk on Ω with jump rates at time s:*

$$\bar{x}_\eta(t-s)a(\xi,\eta)(\bar{x}_\xi(t-s))^{-1}. \tag{0.28}$$

Let for t fixed $(a_s^t)_{s\in[0,t]}$ be the transition probability from time 0 to s, induced by the rates for the transition from ξ to η at time s (for well-posedness see subsection 1(a)).

Define for $s \in [0,t]$ (here with convention $0/0 = 0$):

$$C^t_\xi(s) = h(\bar{x}_\xi(t-s))/\bar{x}_\xi(t-s). \tag{0.29}$$

$$\widehat{\mathcal{G}}_t = \int_0^t \sum_\eta a^t_s(0,\eta) a^t_s(0,\eta) C^t_\eta(s) ds. \tag{0.30}$$

Then consider the two cases:

$$\mathcal{L}(\widehat{\mathcal{G}}_t) \underset{t\to\infty}{\Longrightarrow} \delta_\infty \quad \text{or} \quad \widehat{\mathcal{G}}_t \text{ stochastically bounded.} \tag{0.31}$$

In the first case (write $\underline{u}$ for $(\delta_u)^\Omega$) we find a monotype limit:

$$\mathcal{L}(\widehat{X}(t)) \underset{t\to\infty}{\Longrightarrow} \int_{[0,1]} \widehat{\theta}(du)\delta_{\underline{u}}, \tag{0.32}$$

and in the second case coexistence of types

$$\mathcal{L}(\widehat{X}(t)) \underset{t\to-\infty}{\Longrightarrow} \mu_\theta, \quad \mu_\theta \text{ is nontrivial, i.e. is not mixture of the } \delta_{\underline{u}}. \tag{0.33}$$

If we define P^t to be the law of two independent $(a^t_s)_{s\leq t}$ chains starting in 0 and by N_t the number of returns of the difference process to 0 by time t, then if

$$\inf_{x\geq 0}(h(x)) > 0, \quad P^t(N_t \geq K) \underset{t\to\infty}{\longrightarrow} 1 \quad \forall K \tag{0.34}$$

we find the situation (0.32). $\square$

Remark In the recurrent case one can understand (0.31) better by studying coalescing random walks in a space-time random medium. The question is whether $\widehat{a}$ recurrent implies $\widehat{\mathcal{G}}_t \to \infty$ as $t \to \infty$. We do not have the space to address this problem in the present paper. The background and where $\widehat{\mathcal{G}}_t$ comes from will become clear in Proposition 0.4 in the next subsection and from the duality relation presented in subsection 1(b).

Open problem If $\widehat{a}$ is recurrent then the question is whether it is possible to have $N_t \to \infty$, but such that a rapid decrease of $h(x)$ as $x \to 0$ forces $\widehat{\mathcal{G}}_t$ to be stochastically bounded? This is difficult to resolve since in that case convergence to 0 of the total mass process also slows down. We shall study this question in the hierarchical mean-field limit in a forthcoming paper.

Remark In the recurrent case a second question is the behavior of $\widehat{X}(t)$ in colonies, where the total mass is very large. Do we have monotype clusters? We expect the answer to be always yes. Here is a partial (for technical reasons) answer.

Corollary *Let $h(x) \equiv c$.*
Then if $\mathcal{G}_t$ denotes $\int_0^t \hat{a}_s(0,0)ds$, we can say the following about the rare populated sites:

$$(\mathcal{L}(\widehat{X}(t) | \bar{x}_\xi(t) \geq \varepsilon\mathcal{G}_t)) \Rightarrow \int_{[0,1]} \delta_{\underline{u}}\widehat{\theta}(du) \quad \text{as } t \to \infty. \quad \square \tag{0.35}$$

(c) Representation of the total mass and relative weights process. In this subsection we derive characterizations and properties of important functionals of the process, the total mass per site in Proposition 0.1, its path properties, in particular zeros of the component processes and the behavior near such zeros in Proposition 0.2 and of the relative proportions of types at a site *conditioned* on the complete total mass process in Proposition 0.4. The latter leads us to Fleming-Viot processes evolving in time-space random media, where in certain cases the medium even has *singularities* and these processes are defined via Proposition 0.3.

(i) The dynamics of the total mass process
Define first the total mass analogue of $\mathcal{E}$ (recall (0.6)), namely $\bar{\mathcal{E}}$ by:

$$\bar{\mathcal{E}} = \Big\{ \bar{X} \in (\mathbb{R}^+)^\Omega \Big| \sum_\xi \bar{x}_\xi \gamma(\xi) < \infty \Big\}. \tag{0.36}$$

For the evolution of the total mass in the colonies $\xi \in \Omega$ or the evolution of a single family indexed by $i \in [0,1]$ it follows from Theorem 1 in connection with Ito's formula, that we can use systems of coupled stochastic differential equations to define a version of the total mass functionals of the process. We prove in subsection 2:

Proposition 0.1. *A version of the process $\bar{X}(t) = (\bar{x}_\xi(t))_{\xi\in\Omega}$ defined by $\bar{x}_\xi(t) = x_\xi(t)([0,1])$ is given by the unique strong solution to the following system of stochastic differential equations (recall $g(x) = xh(x)$):*

$$\begin{aligned} d\bar{x}_\xi(t) &= \sum_{\xi'} a(\xi,\xi')(\bar{x}_{\xi'}(t) - \bar{x}_\xi(t))dt + h(\bar{x}_\xi(t))c(\bar{x}_\xi(t))\bar{x}_\xi(t)dt \\ &\quad + \sqrt{(2g(\bar{x}_\xi(t))}dw_\xi(t), \quad \xi \in \Omega. \end{aligned} \tag{0.37}$$

Here $((w_\xi(t))_{t\geq 0})_{\xi\in\Omega}$ are independent standard Brownian motions. The paths are a.s. elements of $C([0,\infty), \bar{\mathcal{E}})$. □

(ii) The zeros of the total mass process
A well-defined relative mass process $\widehat{X}(t)$, can be defined provided that the path properties of the total mass process $\bar{X}(t)$ satisfies certain conditions. In particular these conditions involve the 0's of $\bar{x}_\xi(t)$ and the behavior close to them. In subsection 1(a) that the appropriate path properties of the total mass process are established:

Proposition 0.2. *Assume $c \equiv 0$ and that the initial state of $\bar{X}(t)$ (recall (0.37)) is translation invariant ergodic with strictly positive intensity a.s.. Then the following properties hold.*
(a) *Fix a sequence $\xi_1, \xi_2, \ldots$ of distinct points. Then for $\bar{X}(0) \in \bar{\mathcal{E}}$ with strictly positive limiting mean, there exists a random variable $n_0(T, \bar{X}(0))$, measurable with respect to the Brownian motions up to time T and almost surely finite such that*

$$P(\{\bar{x}_{\xi_1}(t) = \bar{x}_{\xi_2}(t) = \bar{x}_{\xi_n}(t) = 0 \text{ for some } t \leq T, n \geq n_0(T, X(0))\}) = 0, \quad \text{for all } T \in \mathbb{R}^+. \tag{0.38}$$

(b) *If for fixed $\xi \in \Omega$, $\bar{x}_\xi(0) > 0$, then the process $(\bar{x}_\xi(t))_{0\leq t\leq T}$ is either strictly positive or $I_\xi := \{t \in [0,T] : \bar{x}_\xi(t) = 0\}$ is infinite. Both cases can occur depending on the parameters.*
(c) *If $x_\xi(0) > 0, \quad \forall \xi \in \Omega$ and $h(0) = 0$, then*

$$P(\bar{x}_\xi(t) > 0, \quad \forall t \in [0,T], \forall \xi \in \Omega) = 1, \quad \forall\, T \in \mathbb{R}^+. \tag{0.39}$$

(d) *For each ξ, a.s. I_ξ has Lebesgue measure zero and the complement of I_ξ is a countable union of disjoint intervals, called the* excursion intervals*,*
(e)*For each $\xi \in \Omega$,*

$$\text{the local time of } (\bar{x}_\xi(t))_{t\geq 0} \text{ at } 0 \text{ is zero a.s..} \tag{0.40}$$

(f) *Almost surely:*

$$\int_{t_0}^{t_0+\varepsilon} (\bar{x}_\xi(s))^{-1} ds = \infty, \quad \forall \varepsilon > 0 \quad \text{and} \quad t_0 \in I_\xi. \tag{0.41}$$

(g) *Furthermore if $\bar{x}_\xi(t_0) = 0$ and t_0 is a stopping time or the starting time of an excursion away from 0, then there exists an η with $a(\xi, \eta) > 0$, such that:*

$$\int_{t_0}^{t_0+\varepsilon} (\bar{x}_\eta(s)/\bar{x}_\xi(s))ds = +\infty \text{ for all } \varepsilon > 0. \tag{0.42}$$

(h) *For almost every realization of the total mass process, the random walk starting in ξ with $\bar{x}_\xi(t) > 0$ and with rates a_s^t for a jump from ξ to η at time $s \leq t$, given by:*

$$a_s^t(\xi, \eta) = (\bar{x}_\xi(t-s))^{-1} a(\xi, \eta) \bar{x}_\eta(t-s), \tag{0.43}$$

is well-defined. Furthermore, a.s. the random walk does not hit a singularity in finite time, that is, a space-time point at which a zero of the total mass process occurs. □

Remark Zeros of the total mass processes $\bar{x}_\xi(t)$ occur only if the process is qualitatively like a classical branching diffusion, that is, $(h(0) > 0)$ and the immigration is less than a critical value.. However this case is a very important one.

(iii) Preparation: Fleming-Viot processes in a singular random medium

Our next goal is to give a simpler characterization of the process $\widehat{X}(t)$ if the complete total mass process $(\bar{X}(t))_{t \geq 0}$ is given. In order to formulate a result to this effect in the next paragraph (iv) we need to consider first a new class of interacting measure-valued diffusions which is a generalization of interacting Fleming-Viot processes which are characterized in Proposition 0.3. These arise as diffusion limits from a resampling model (in the case of only two or finitely many types, the process is called a system of interacting Fisher-Wright diffusions). In fact we need such processes evolving in randomly fluctuating media. i.e. *interacting Fleming-Viot processes with randomly fluctuating resampling rates.* Worse than that, we have to deal with media with *singularities*, such that singular values at a component may have accumulation points.

In order to define the model, we first discuss the regular case and we begin by recalling the classical (constant) medium case and simply consider a system of *interacting Fleming-Viot* processes with $[0, 1]$ as type space. This is a Markov process with state space $(\mathcal{P}([0, 1]))^\Omega$ where $\mathcal{P}$ stands for probability measures. This process has been defined and analyzed in [DGV, 95]. We define the process rigorously via a *martingale problem.* The key ingredient is the *resampling rate* of a pair of type u and v (u and v from the type space $[0, 1]$) if the colony is in the state x, denoted $H(u, v; x)$. Examples will be given below.

Define the algebra $\widehat{\mathcal{A}}$ of test functions on $(\mathcal{P}([0, 1]))^\Omega$ generated by functions of the form

$$\widehat{F}(\widehat{X}) = \int f(u_1, \dots, u_n) \widehat{x}_{\xi_1}(du_1) \dots \widehat{x}_{\xi_n}(du_n), \tag{0.44}$$

where $f \in C([0, 1]^n)$. Note that this function can be continued to $(\mathcal{M}([0, 1]))^\Omega$, but it is only bounded on $(\mathcal{P}([0, 1]))^\Omega$. Accordingly we define the differential operators $\frac{\partial F}{\partial x_\xi}, \frac{\partial F}{\partial x_\xi \partial x_\eta}$ as in (0.15).

The generator $\widehat{G}$ of the martingale problem associated with the system of (generalized) interacting Fleming-Viot processes is:

$$(0.45)\qquad (\widehat{G}\widehat{F})(\widehat{X}) = \sum_{\xi}\Bigg[\int_0^1 \frac{\partial \widehat{F}}{\partial \widehat{x}_\xi}(\widehat{X})(u)\Big\{\sum_{\xi'} a(\xi,\xi')(\widehat{x}_{\xi'}(du)-\widehat{x}_\xi(du))\Big\}$$
$$+\Big\{\int_0^1\int_0^1 \frac{\partial^2 \widehat{F}}{\partial \widehat{x}_\xi \partial \widehat{x}_\xi}(\widehat{X})(u,v)\left[\int_0^1 H(u,w;\widehat{x}_\xi)\widehat{x}_\xi(dw)\right]\widehat{x}_\xi(du)\delta_u(dv)$$
$$-\int_0^1 \frac{\partial^2 \widehat{F}}{\partial \widehat{x}_\xi \partial \widehat{x}_\xi}(\widehat{X})H(u,v;\widehat{x}_\xi)\widehat{x}_\xi(du)\widehat{x}_\xi(dv)\Big\}\Bigg].$$

The case of systems of classical interacting Fleming-Viot processes is obtained for $H(u,v;\widehat{x}_\xi)\equiv\delta$ respectively the weighted sampling case by $H(u,v,x_\xi)=H(u,v)$. To be more concrete consider $H(u,v,x_\xi)\equiv\delta$. In this case in the second term the third integral is independent of u,v and x_ξ and hence equals δ.

If in the classical case we consider the evolution of a single type i say, the law of the weight $\widehat{x}^i_\xi$ of that type in colony ξ is given by the following system of stochastic differential equations

$$(0.46)\qquad d\widehat{x}^i_\xi(t)=\sum_{\xi'} a(\xi,\xi')(\widehat{x}^i_{\xi'}(t)-\widehat{x}^i_\xi(t))dt+\sqrt{2\delta\widehat{x}^i_\xi(t)(1-\widehat{x}^i_\xi(t))}\,dw_\xi(t),\quad \xi\in\Omega,$$

which is a system of interacting *Fisher-Wright diffusions*.

Definition 0.3. We call the function $H(u,v,x)$ the *resampling rate* of the pair of types (u,v) in state x and $a(\xi,\xi')$ the *migration rate* from ξ' to ξ. $\square$

In order to describe the process $\widehat{X}(t)$ of relative weights it is very important to deal with a *time-inhomogeneous* version of the above process allowing *singular rates*. We consider H and $a(\cdot,\cdot)$ depending both on t and $\xi\in\Omega$ and being of the form (use the convention $0\infty=0$).

$$(0.47)\qquad H_{t,\xi}(u,v;\widehat{z})=d_\xi(t),\quad a^t(\xi,\eta)=m_\xi(t)a(\xi,\eta)m^{-1}_\eta(t).$$

If $d_\xi(t)$ is finite and $m_\xi(t)$ finite and never 0 and both functions are continuous then we could define the needed process with resampling rates $d_\xi(t)$ and migration rates $a^t(\xi,\eta)$ in the classical way via a martingale problem associated with $\widehat{G}$ in (0.45) (compare Proposition 0.4 below).

However we shall need to allow for the possibility that both $m_\xi(t)$ and $d_\xi(t)$ take the value $+\infty$. Namely if we apply the construction above to the relative weight process $\widehat{X}(t)$, we shall see later that we need $m_\xi(t)=(\bar{x}_\xi(t))^{-1}$ and $d_\xi(t)$ containing the same factor. Hence a problem occurs due to the fact that the path of $x_\xi(t)([0,1])$ might have zeros, recall Proposition 0.2 here. In order to define things properly in this context we need some preparation.

We must now impose a condition on the collection, m_ξ, namely that they are admissible. To define this consider the random walk starting at ξ with $m_\xi(0)<\infty$ and with transition rates $a^t(\xi,\eta)$ given by 0.47 at times t for which $m_\xi(\cdot)$ has no singularities and which leaves ξ instantaneously if $m_\xi(t)=+\infty$ and $m_\eta(t)>0$ for some η with $a(\xi,\eta)>0$. In this case it jumps to a point η chosen with probability proportional to $m_\eta(t)a(\xi,\eta)$. Then the rates

$$(0.48)\qquad \{m_\xi(t),\ t\geq 0,\ \xi\in\Omega\},\quad \infty\geq m_\xi(t)>0\quad \text{for all } \xi\in\Omega,$$

are said to be *admissible* if a.s. this random walk does not escape to infinity, that is leave every finite set, in finite time.

Altogether we require for the two collections $\{(d_\xi(t))_{t\geq 0}\}_{\xi\in\Omega}$ and $\{(m_\xi(t))_{t\geq 0}\}_{\xi\in\Omega}$, that they have components in $C([0,\infty),\ \bar{\mathbb{R}}^+)$ which satisfy:

(0.49)
- (0) They take the value $+\infty$ only in a set of points of Lebesgue measure 0.
- (i) The value $+\infty$ never occurs simultaneously for countably many sites.
- (ii) If m_ξ and d_ξ are $+\infty$ at the same time t_0 we assume that $m_\xi(t)/d_\xi(t)$ converges to a finite limit as $t\to t_0$.
- (iii) The rates $(m_\xi)_{\xi\in\Omega}$ are admissible.
- (iv) The singularities of m_ξ and d_ξ are not integrable from either side.
- (v) $m_\xi(t)/m_\eta(t)$ is not integrable at singularities t of m_ξ, such that m_ξ is finite for some interval to the left of t (here $\eta\neq\xi$) and η satisfies $a(\xi,\eta)>0$.

We then define the following martingale problem for a given collection $\{(m_\xi,d_\xi),\xi\in\Omega\}$, which specifies the local characteristics between singularities:

Definition 0.4. (Martingale problem with singularities in the rates).
For all $\widehat{F}\in\widehat{\mathcal{A}}$ (recall (0.44)) the following should hold for the solution of the martingale problem w.r.t. $\widehat{\mathcal{A}}$.

Let $\widehat{F}$ depend on components $x_{\xi_1},\dots,x_{\xi_n}$. Consider times u such that for the given $\widehat{F}$ no singularities occur for $m_\xi(r),d_\xi(r)$ and $\xi\in\{\xi_1,\dots,\xi_n\}$. Define for such values of u the value of the operation $\widehat{G}_r(\widehat{F})$ for $\widehat{F}\in\widehat{\mathcal{A}}$ by replacing in the r.h.s. of (0.45) $a(\cdot,\cdot)$ by $a^r(\cdot,\cdot)$ and H by $H_{r,\xi}$, with $a^r,H_{r,\xi}$ as in (0.47).

Next define locally good time intervals beginning in s by:

$$h_t(s,\widehat{F})=1\!\!\mathrm{I}\,\{(d_{\xi_i}(r)<\infty,\quad m_{\xi_i}(r)<\infty,\qquad \forall r\in(s,t),\ i=1,\dots,n)\}\,. \tag{0.50}$$

Then the $(\widehat{G}_t,\delta_{\widehat{X}})$-*martingale problem* w.r.t. $\widehat{\mathcal{A}}$ asks for a law on a subset of paths with values in $(\mathcal{P}([0,1]))^\Omega$ (see (iii) below for this regularity requirement) such that for the canonical process the following holds:

(i) Let $s\geq 0$ be any number such that $d_{\xi_i}(s)+m_{\xi_i}(s)<\infty, i=1,\dots,n$ and let $T'=\inf(t>s|h_t(s,\widehat{F})=0)$. Then for $T<T'$ the process below is a continuous martingale w.r.t. the canonical filtration:

$$\left(\left\{\widehat{F}(\widehat{X}_{t\wedge T})-\widehat{F}(\widehat{X}_s)-\int_s^{t\wedge T}\widehat{G}_u(\widehat{F})(\widehat{X}_u)du\right\}\right)_{t\geq s}, \tag{0.51}$$

and

$$\widehat{X}_0=(\widehat{x}_\xi(0))_{\xi\in\Omega}\quad\text{with } \widehat{x}_\xi(0)=\widehat{x}_\xi\quad \forall\xi \text{ with } m_\xi(0)+d_\xi(0)<\infty. \tag{0.52}$$

(ii) Define the $\{(d_\xi(t))_{t\geq 0},\ (m_\xi(t))_{t\geq 0},\ \xi\in\Omega\}$-dependent set of time points I:

$$I=\{t|\ d_{\xi_i}(t) \text{ or } \ m_{\xi_i}(t)=+\infty \qquad \text{for some } \ i\in\{1,\dots,n\}\}\,. \tag{0.53}$$

We require that the following holds for the closure of the set of all $T\in I$ such that either $(T,T+h)$ or $(T-h,T)$ have no intersection with I, for some $h>0$, and there exists for those ξ_i with $m_{\xi_i}(t)=+\infty$ a site η with i $a(\xi_i,\eta)>0$ and $m_\eta(t)<\infty$:

$$\mathcal{L}(\widehat{F}(\widehat{X}_T))=w-\lim_{h\to 0}\mathcal{L}(\widehat{F}(\widehat{X}_{T\pm h})). \tag{0.54}$$

For all other $T\in I$ and $i\in\{1,\dots,n\}$ we require $\hat{x}_T(\xi)=\nu$, where ν is a fixed arbitrary element of $\mathcal{P}([0,1])$.

(iii) Define for every $\xi\in\Omega$ the set $\mathcal{T}_\xi=\{t\in[0,\infty)|m_\xi(t)+d_\xi(t)=+\infty\}$. Then we require that for every $\xi\in\Omega$:

$$(\hat{x}_\xi(t))_{t\geq 0} \text{ is continuous on the complement of } \mathcal{T}_\xi. \qquad\square \tag{0.55}$$

Remark We could use a simpler convention instead of (0.54), namely pick a $\nu \in \mathcal{P}([0,1])$ and require:

$$\hat{x}_\xi(t) = \nu \qquad \forall \xi \ \text{ with } m_\xi(t) = +\infty. \tag{0.56}$$

However then we do not have continuity properties of finite dimensional distribution, where we could have it. The condition before (0.54) is necessary, if a point ξ with a singularity of m at t has the property that all η with $a(\xi,\eta) > 0$ exhibit also a singularity of m at t. In this case in general the law of $\hat{x}_\xi(t)$ will not have a continuity point at t, since we will not have convergence from either side as we approach t.

Proposition 0.3. *Assume $H(u,v;\widehat{x}_\xi(t)) = d_\xi(t)$ and let a^t be as in (0.47). Assume furthermore that $X \in (\mathcal{P}(\mathbb{N}))^\Omega$ and the collections $(d_\xi(t))_{\xi\in\Omega}$ and $(m_\xi(t))_{\xi\in\Omega}$ with $d_\xi, m_\xi \in C([0,\infty),\bar{\mathbb{R}}^+)$ satisfy (0.49). Then the $(\widehat{G}_t, \delta_{\widehat{X}})$-martingale problem w.r.t $\widehat{\mathcal{A}}$ is well-posed.* $\square$

The proof of these results is in subsection 1(b) step 3.

(iv) *The relative weights process conditioned on $\bar{X}$*

Crucial for both the description and the analysis of our model is the following description of the dynamics of the relative weights process conditioned on $(\bar{X}(t))_{t\geq 0}$ which generalizes a result of Perkins [P2] and which relates branching to resampling models. The basis is the above Proposition 0.3 in connection with Proposition 0.2.

First we have to deal with the fact that relative weights in the case of 0 mass have been set equal to the 0-measure in (0.22) which is not natural from the point of view of the state space of the process which are probability measures. We therefore have to adapt our convention $\hat{x}_\xi(t) =$ 0-measure for t with $\bar{x}_\xi(t) = 0$ by using a cemetery type $\tilde{\Delta}$ and putting $\hat{x}_\xi(t) = \delta_{\tilde{\Delta}}$ in that case.

Definition 0.5. We say that $(\widehat{X}_t)_{t\geq 0}$ is an *admissible modification* of the given relative weights process of $(X_t)_{t\geq 0}$ if both processes agree in all components except possibly where the process $(X_t)_{t\geq 0}$ has a component which is the 0 measure. $\square$

The dynamic of the relative weights process $\widehat{X}(t)$ conditioned on the total mass process has a nice description, namely (the proof is in subsection 2):

Proposition 0.4. *Assume that $c \equiv 0$. The process $\widehat{X}(t)$ conditioned on the total mass process $(\bar{X}(t))_{t\geq 0}$ has an admissible modification, which is a time-inhomogeneous Fleming-Viot process with resampling rate at site ξ and at time t given by*

$$h(\bar{x}_\xi(t))(\bar{x}_\xi(t))^{-1} \tag{0.57}$$

and migration rate from ξ' to ξ given at time t by

$$a(\xi,\xi')\bar{x}_{\xi'}(t)/\bar{x}_\xi(t). \qquad \square \tag{0.58}$$

Remark The quantities in (0.57) and (0.58) are exactly the ones giving $(a^t_s)_{s\in[0,t]}$ and $C^t_\xi(s)$ in Theorem 1(d) and arise from using the dual process.

Remark Near the points t, where $\bar{x}_\xi(t) = 0$, the law of $\hat{x}_\xi(s)$ looks like a Fleming-Viot process with immigration-emigration in equilibrium where the immigration source is $\sum_\eta a(\xi,\eta)x_\eta(t)$. This holds provided the latter quantity is positive.

(d) The historical process: Construction and longtime behavior. In branching systems the description of the population at time t does not exhaust the information of interest. One would like to know in addition, which families contribute to a given population at time t or where the family was originally located and how are the relatives spread out in space at different times $s \leq t$. Another question is to determine the family decomposition of a population which started to evolve at time s and is observed at time 0 and with $s \to -\infty$, a so-called infinitely old population. In order to incorporate this information one considers the so-called historical process.

The main purpose of this subsection is to first construct the historical process as the unique solution of a martingale problem in Theorem 3 and then secondly to study its behavior as $t \to \infty$ in Theorem 4. This provides the tools to give a family decomposition of an infinitely old population.

(i) Basic ideas and construction of the historical process via a martingale problem

We first explain the basic idea. Return for the moment to the particle picture. There the historical process arises by associating with every individual alive at time t in a branching random walk starting at time $s < t$ its *type* and its *path of descent*. This will then complement and further refine the information already contained in our process, namely which types make up a given colony. Hence we consider paths with a mark, namely paths with values in $\Omega \times [0,1]$ the first describing location the second the type. Recall that the number of types actually present in the system is for $t > 0$ a.s. *countable*. Note that in our model the type will not change along a path. The evolution of a path of descent is called the *path process*, which is a time inhomogeneous Markov process.

Note that for a population of age t the path of descent does not characterize the individual. However since we later consider the equilibrium case of an infinitely old population (when the path does characterize the individual), we suppress at this point the inclusion of the family label in the description of the process to simplify notation. Thus we can represent a configuration of paths of descent in the particle picture by a *measure on path space* (path with values in $\Omega \times [0,1]$ and time running from s to t), by associating with every path a δ-measure on that path and then form the sum over all particles, present at time t. This gives us for fixed t an integer-valued and locally finite measure on path space. Hence the historical process will be a measure valued process, which is induced by a *path-valued process* describing the evolution of ancestral path. Note that the measure-valued process is still a *Markov process*. In order to have for every time t a path of the same length we continue the path beyond the time t considered, as well as to the left of s, in both cases as a constant path.

In addition we want to consider this construction for the process $X(t)$ in equilibrium and started long time ago. Mathematically, this then amounts to constructing the path valued process for the equilibrium. We do this by starting the process at time $-\infty$ (i.e. $s \to -\infty$).

Next we pass again from the particle model to the continuum mass limit. If we now give every path mass ε, increase the initial intensity of individuals as ε^{-1}, speed up the branching rate per particle by ε^{-1} and scaling the branching rate at $\lfloor x\varepsilon \rfloor$ to converge to $h(x)$ we get for $\varepsilon \to 0$ a limit. In this *diffusion limit* we then get a process with values in the locally finite measures on paths. In the diffusion limit the historical process is still Markov and we are able to describe it via a *martingale problem*, which we now turn to below.

Finally it is important to remark that in the diffusion limit the branching points become "identifiable" from the measure on the path because two particles whose past trajectories are identical over a time interval in the past must be descendants coming from a single common ancestor and the time at which they branched is precisely the time at which the common trajectory ends. This is in contrast to the particle case in which two particles with identical trajectories over an infinite time interval in the past must have a common ancestor but may have branched at a time earlier than the end of their common trajectory.

In order to carry out the ideas presented above on a rigorous mathematical level we first formulate the corresponding martingale problem. For this purpose we need some ingredients and notations which we introduce next.

The historical process will be denoted by $(X^*(t))_{t\geq 0}$ and the *state space* $\mathcal{E}^*$ will be a subset (recall the introduction of $\mathcal{E}$ for the state space of $X(t)$ in (0.6)) of:

$$\mathcal{M}(D(\mathbb{R}, \Omega \times [0,1])). \tag{0.59}$$

The set $\mathcal{E}^*$ we define by, here y abbreviates an element of $D(\mathbb{R}, \Omega \times [0,1])$:

$$\mathcal{E}^* = \{\mu \in \mathcal{M}\{D(\mathbb{R}, \Omega \times [0,1])\}| \quad (\mu(\{y|y(t) \in \{\xi\} \times [0,1]))_{\xi \in \Omega} \in \bar{\mathcal{E}}, \quad \forall t \in \mathbb{R}\}. \tag{0.60}$$

The next step is to define a martingale problem on the space $\mathcal{E}^*$, which describes the historical process based on the path process describing the evolution of the paths of descent. Note that the historical process will be a *time inhomogeneous process* as is the path-process. However passing with the latter to the *time-space process* we obtain a time-homogeneous generator for the path process.

The first ingredient which we shall need is the *generator* $\widetilde{A}$ *of the (time-space) path process*, which acts on functions

$$\Phi(s,y) \quad \text{with} \quad s \in \mathbb{R},\ y \in D(\mathbb{R}, \Omega \times [0,1]). \tag{0.61}$$

This generator of the path process will be specified in terms of the generator of the underlying flow process on $\Omega \times [0,1]$, which is denoted by A and is given by:

$$(Af)(\xi,u) = \sum_{\xi'} (a(\xi',\xi) - \delta(\xi,\xi'))f(\xi',u), \quad f \in L_\infty(\Omega \times [0,1]). \tag{0.62}$$

To specify $\widetilde{A}$ in terms of A consider functions $g_j : \mathbb{R} \times (\Omega \times [0,1]) \to \mathbb{R}$ which are bounded and have in the first variable bounded continuous derivatives. Then form for a given collection $t_1 < t_2 < \cdots < t_n,\ n \in \mathbb{N}$ the function:

$$\Phi(t,y) = \prod_{j=1}^{n} g_j(t, y(t \wedge t_j)). \tag{0.63}$$

The algebra generated by these functions is denoted $\widetilde{\mathcal{A}}$, which is measure determining on $D(\mathbb{R}, \Omega \times [0,1])$ (see [D]).

Define the action of $\widetilde{A}$ on Φ as follows. For t such that $t_k < t \le t_{k+1}$ we set:

$$\begin{aligned}(\widetilde{A}\Phi)(t,y) \quad &= \quad \prod_{j=1}^{k} g_j(t,y(t\wedge t_j)) \left[\left(\frac{\partial}{\partial t} + A \right) \prod_{j=k+1}^{n} g_j(t,y(t \wedge t_j)) \right] \\ &+ \quad \left[\frac{\partial}{\partial t} \prod_{j=1}^{k} g_j(t,y(t\wedge t_j)) \right] \left[\prod_{j=k+1}^{n} g_j(t,y(t\wedge t_j)) \right].\end{aligned} \tag{0.64}$$

Another important ingredient to define the generator of X^* will be the stopped path

$$y^r \in D(\mathbb{R}, \Omega \times [0,1]) \quad \text{defined by } y^r(t) = y(t \wedge r) \tag{0.65}$$

and the corresponding measure $X^{*,r}$, which is the image of X^* induced by the map $y \to y^r$ hence:

$$X^{*,r} \in \mathcal{M}(D(\mathbb{R}, \Omega \times [0,1])). \tag{0.66}$$

Furthermore we need the projection of $(y(t))_{t \ge 0}$ on its space component Ω, which we denote by:

$$\pi_1 y, \ \text{i.e. } (\pi_1 y)(t) \in \Omega. \tag{0.67}$$

Now we can write down the *historical martingale problem* for given starting point X_0^*:

Definition 0.6. Let $X_0^* \in \mathcal{E}^*$ and a starting time $s \in \mathbb{R}$ be given. For every function Φ of the form given in (0.63) and for every t, s with $t \geq s$ the process $(X^*(t))_{t\geq 0}$ satisfies the three conditions:

$$\begin{array}{ll} (i) & X^*(s) = X_0^*, \\ (ii) & \left\{ \langle X^*(t), \Phi(t,\cdot)\rangle - \langle X^*(s), \Phi(s,\cdot)\rangle - \int_s^t \langle X^{*,r}(r), (\widetilde{A}\Phi)(r,\cdot)\rangle dr \right\}_{t\geq s}, \\ & \text{is a martingale,} \\ (iii) & \text{the corresponding increasing process is} \\ & \left(\int_s^t \left\{ \int_{D(\mathbb{R},\Omega\times[0,1])} [\Phi^2(r,y)h(X^{*,r}(\{\widetilde{y}|\pi_1\widetilde{y}(r) = \pi_1 y(r)\}))](X^{*,r}(r))(dy) \right\} dr \right)_{t\geq s}. \quad \square \end{array} \tag{0.68}$$

The final step in order to define the historical process is to prove:

Theorem 3. *The martingale problem (0.68) on $[s,\infty)$ with $s \in \mathbb{R}$ is well-posed for $X^*(s) = X_0^* \in \mathcal{E}^*$ and the solution is a law on $C([s,\infty), \mathcal{E}^*)$.* $\square$

Definition 0.7. **(a)** The canonical process corresponding to the unique solution of the martingale problem (0.68), $(X^*(t))_{t\geq s}$, is called the *historical process*, with starting time s and initial point X_0^*.

(b) Similarly we define a new process as the image arising by projecting the set $\Omega\times[0,1]$ (in (0.59)) just onto Ω. This process we will call $\bar{X}^*(t)$, the *historical total mass process* and this process is defined by

$$\bar{X}^*(t)(A) = X^*(t)(\{y|\pi_1 y \in A\}), \qquad A \in \mathcal{B}(D(\mathbb{R},\Omega)), \tag{0.69}$$

where π_1 projects onto the first element of $\Omega\times[0,1]$. $\square$

(ii) The historical process of an infinitely old population and its family decomposition

The next point is to construct the historical process for an infinitely old population, that is to let in the Theorem 3 the starting time s of our system tend to $-\infty$. This works of course only in the stable context, where the original system $X(t)$ approaches as $t\to\infty$ a nontrivial equilibrium state, that is $\widehat{a}$ has to be transient.

Theorem 4. *Assume that the kernel $\widehat{a}$ is transient. Let $\mathcal{L}(Y^*)$ be concentrated on constant path, translation invariant and ergodic with $E(Y^*(\{y|y(0)\in\{\xi\}\times[0,1]\})) = \bar{\theta}^* \in [0,\infty)$ and $E(Y^*(\{y|y(0)\in\{\xi\}\times\cdot\})) = \theta^* \in \mathcal{M}([0,1])$.*

Then there exist limit states, $\widetilde{\nu}_{\theta^*}^*$ and $\widetilde{\nu}_{\bar{\theta}^*}^*$, depending on Y^* only through θ^* respectively $\bar{\theta}^*$ such that:

$$\mathcal{L}(X^*(t)|X^*(s) = Y^*) \underset{s\to-\infty}{\Longrightarrow} \widetilde{\nu}_{\theta^*}^*(t), \tag{0.70}$$

$$\mathcal{L}(\bar{X}^*(t)|X^*(s) = Y^*) \underset{s\to-\infty}{\Longrightarrow} \widetilde{\nu}_{\bar{\theta}^*}^*(t). \qquad \square \tag{0.71}$$

We can now consider in (0.70) respectively (0.71) for our purposes without loss of generality $t = 0$ and we obtain probability measures

$$\nu_{\theta^*}^* \in \mathcal{P}[\mathcal{M}(D(\mathbb{R},\Omega\times[0,1]))], \qquad \nu_{\bar{\theta}^*}^* \in \mathcal{P}\left[\mathcal{M}(D(\mathbb{R},\Omega))\right]. \tag{0.72}$$

Definition 0.8. We call a realization of $\nu_{\theta^*}^*$ the *equilibrium historical process* $\widetilde{X}^*$, which is a random measure with values in $\mathcal{M}(D(\mathbb{R},\Omega\times[0,1]))$. Similarly a realization of $\nu_{\bar{\theta}^*}^*$ is called *equilibrium historical total mass process.* $\square$

Let us mention also the connection between $\nu_{\theta^*}^*$ and ν_θ of (0.26). Define $\widetilde{\widetilde{X}}^* = (\widetilde{\widetilde{x}}_\xi^*)_{\xi\in\Omega}$ by

$$\widetilde{\widetilde{x}}_\xi^*(A) = \widetilde{X}^*(\{y|y \text{ has type } i \in A \text{ and } y(0) = \xi\}), \qquad \text{for all } A \in \mathcal{B}([0,1]),\ \xi\in\Omega, \tag{0.73}$$

where $\widetilde{X}^*$ is distributed according to $\nu_{\theta^*}^*$.

Then $\mathcal{L}(\overset{\approx}{X}{}^*))$ and ν_θ from Theorem 2 are related as follows:

Corollary *Choose $\theta^* = \theta \in \mathcal{M}([0,1])$. Then:*

$$\mathcal{L}(\overset{\approx}{X}{}^*) = \nu_\theta. \qquad \square \tag{0.74}$$

We can use now the equilibrium historical process to define a *family decomposition*. Introduce the following equivalence classes on $D(\mathbb{R}, \Omega \times [0,1])$ and $D(\mathbb{R}, \Omega)$. Here π_1, and π_2 denote projections on $D(\mathbb{R}, \Omega)$ and $D(\mathbb{R}, [0,1])$.

For $y_1, y_2 \in D(\mathbb{R}, \Omega)$ define:

$$y_1 \sim y_2 \quad \text{iff } \exists T : y_1(s) = y_2(s) \quad \forall s \leq T. \tag{0.75}$$

For $y_1', y_2' \in D(\mathbb{R}, \Omega \times [0,1])$ define:

$$y_1' \approx y_2' \quad \text{iff} \quad \exists T : (\pi_1 y_1')(s) = (\pi_1 y_2')(s) \quad \forall s \leq T \quad \text{and } \pi_2 y_1' = \pi_2 y_2'. \tag{0.76}$$

Then we get, that $\widetilde{X}^* \in \mathcal{M}(D(\mathbb{R}, \Omega \times [0,1]))$ can be decomposed as follows

$$\widetilde{X}^* = \sum_{i \in \mathfrak{I}} \widetilde{X}^*|_{E_i}, \tag{0.77}$$

where E_i is a $\sim$ or $\approx$ equivalence class and $\mathfrak{I}$ stands for $\mathfrak{I}^\sim, \mathfrak{I}^\approx$ which is the set of all equivalence classes with positive mass with respect to $\sim$ and $\approx$.

Proposition 0.5. *Assume that $\hat{a}$ is transient and θ^* a finite, nonzero measure on* $[0,1]$.

(a) *Under $\nu_{\theta*}^*$ there are w.r.t. to both $\sim$ or $\approx$ a.s. countably many families with positive mass.*
(b) *Let $\mathfrak{I}^\sim$ and $\mathfrak{I}^\approx$ be the random set of equivalence classes with respect to $\sim$ and $\approx$. Then*

$$\mathfrak{I}^\sim = \mathfrak{I}^\approx \quad \nu_{\theta^*}^* - a.s. \quad \square \tag{0.78}$$

Remark: Due to the fact that the migration mechanism and the branching mechanism are independent of the type, we can get a version of the equilibrium historical process corresponding to the intensity measure θ. We take the cluster decomposition of the process corresponding to the process with total mass parameter $\bar{\theta}$. Then assign independently a random type to each family cluster drawn at random according to the distribution $\hat{\theta}$.

1. Path properties of the total mass and relative weights process

In this chapter we derive some basic information about the finer path properties of the total mass and the relative weight processes needed to prove Proposition 0.2 and Proposition 0.4 and which is used in section 2 in a crucial way for the proofs of the uniqueness result. In particular we shall construct for interacting Fleming-Viot processes with singular rates a dual process that will be used to establish that the process of relative weights conditioned on the total mass process is well-defined by the martingale problem specified in chapter 0(c).

It will turn out later in chapter 2 that zeros of the total mass process create singularities for the rates of the relative weights process. Furthermore the dual process which is constructed in section 1(b) is a system of coalescing random walks in a randomly fluctuating environment the latter formed by the total mass process. For this reason the properties of the paths $(\bar{x}_\xi(t))_{t \geq 0}$, in particular the structure of the *set of its zeros* as well as the behavior close to these zeros and simultaneous zeros at different sites play an important role in the sequel. This is considered in section 1(a).

In subsection 1(b) we shall collect facts on random walks and coalescing random walks with singular rates which are needed to prove the duality. The duality is constructed in subsection 1(c).

(a) The zeros of the total mass process (Proof Proposition 0.2). The total mass process is given by the unique strong solution of the system of stochastic differential equations of Proposition 0.1. In this subsection we identify the properties of this process that are needed to establish uniqueness properties of the relative weights process needed to prove Theorem 1. In particular we obtain criteria under which the total mass process is strictly positive and we analyze the structure of its zeros and the behavior near zero when it is only non-negative. This will prove Proposition 0.2 from subsection 0(c) except part (h), which we shall prove in subsection 1(b).

A key point to realize is that the zeros of a component process $(\bar{x}_\xi(t))_{t\geq 0}$ are of two kinds. There is a set of at most countably many zeros which are starting or endpoints of an excursion of the process (into the strictly positive reals) and then there are other zeros (typically uncountably many if there are zeros at all) which are not of that type but occur as limit points of the 'excursion zeros'. Those zeros where excursions start are going to be of particular relevance for us and the set of those zeros is always countable.

(i) Preliminaries (Zeros and behavior near zeros for the mean-field case)
In this step we derive some preparatory results needed for our argument proving Proposition 0.2 in the sequel. As a key step we first consider a simpler mean-field equation, whose one-dimensional components are independent versions of the following $\mathbb{R}^+$-valued diffusion $((Z(t))_{t\geq 0}$ defined as the strong solution of:

$$dZ(t) = c(\theta(t) - Z(t))dt + \sqrt{2g(Z(t))}dw(t). \tag{1.1}$$

Here $c \geq 0$ and $\theta(t)$ is a non-negative continuous function and g is a non-negative locally Lipschitz continuous function, $g(z) = zh(z)$. Note that $\sqrt{2g(z)}$ satisfies the Yamada-Watanabe condition so that (1.1) has a unique non-negative strong solution and the Ikeda-Watanabe comparison theorem ([RW], Theorem 43.1) for different initial conditions and $\theta's$ is applicable. The existence of zeros of (1.6) depends on

$$d = \lim_{x\downarrow 0} h(x). \tag{1.2}$$

Note that by Girsanov the existence of zeros for (1.1) is equivalent to the existence of zeros of

$$d\widehat{Z}(t) = c\theta(t)dt + \sqrt{2g(\widehat{Z}(t))}dw(t). \tag{1.3}$$

For Lemmas (1.1)-(1.4) below we assume that

$$\theta(t) \equiv \theta > 0.$$

Lemma 1.1. *Let $\{Z(t)\}_{t\geq 0}$ be the unique strong solution of (1.1) with initial condition $Z(0) = z_0 > 0$ and $\theta(\cdot) \equiv \theta \geq 0$. Define $T_0 = \inf(t|Z(t) = 0)$. If $\theta + d > 0$, then we have the dichotomy:*

$$if\ \frac{c\theta}{d} \geq 1, then \quad T_0 = +\infty \quad a.s. \tag{1.4}$$

$$if\ \frac{c\theta}{d} < 1, then \quad T_0 < +\infty \quad a.s. \qquad \square \tag{1.5}$$

Remark In particular if $h(0) = 0$, $Z(0) > 0$, then the process $Z(t)$ never hits zero.

Proof We first consider the process $\widehat{Z}(t)$ and let $\widehat{T}_0 = \inf\{t|\widehat{Z}(t) = 0\}$. By the classical diffusion theory, the behavior of the $\widehat{Z}(t)$ depends on the scale function $s(\cdot)$ and speed measure $m(dx)$ (see [RW, 87], V(48,6)). The natural scale $s(\cdot)$ and speed measure $m(dx)$ can be explicitly calculated, namely with $x(y) = s^{-1}(y)$:

$$\frac{ds(v)}{dv} = \exp\left[-\int_\theta^v \frac{c\theta}{wh(w)}dw\right], \tag{1.6}$$

$$s(v) = \int_0^v \exp\left[-\int_\theta^y \frac{c\theta}{wh(w)}dw\right]dy, \tag{1.7}$$

$$m(dy) = (q(y))^{-2}dy, \quad q(y) := (s'(x(y))\sqrt{2x(y)h(x(y))}). \tag{1.8}$$

We first consider the case $\frac{c\theta}{d} = 1$. Since we have assumed that h is locally Lipschitz we can apply the Ikeda-Watanabe comparison theorem and use $h(x) = d + dbx$ as a comparison function. Then the scale function for $\widehat{Z}$ is given by

$$s'(x) = \exp\left(-\int_\theta^x \tfrac{1}{w(1+bw)}dw\right) \quad \text{and}$$
$$s(x) = const \cdot (\log x + bx)$$

and we can take the *const* equal to 1 without loss of generality.

Note that 0 maps to $-\infty$ so we need to prove that $-\infty$ is inaccessible for the process in its natural scale, $s(\widehat{Z}_t)$. For $-\infty < y < 0$ we have:

$$q(y) = \frac{(1 + bx(y))^{\frac{3}{2}}}{x(y)^{\frac{1}{2}}}. \tag{1.9}$$

Note that

$$\lim_{y\to-\infty} \frac{s^{-1}(y)}{e^y} = \lim_{x\to 0} \frac{x}{xe^{bx}} = 1 \tag{1.10}$$

and therefore

$$q(y) \sim \frac{(1 + be^y)^{3/2}}{e^{y/2}}. \tag{1.11}$$

Then (cf. RW V.45),

$$m(dy) = \frac{e^y}{(1 + be^y)^3}dy \tag{1.12}$$

and we have for any $r \in \mathbb{R}$:

$$\int_{-\infty}^r |y|m(dy) < \infty. \tag{1.13}$$

Therefore by the standard criterion $-\infty$ is an entrance boundary (cf. RW V.51.3) which implies that $\widehat{T}_0 = +\infty$, a.s. when $\frac{c\theta}{d} = 1$. The same result for $\frac{c\theta}{d} > 1$ immediately follows from the Ikeda-Watanabe comparison theorem. We then obtain (1.4) by Girsanov.

We now consider the case $0 < \frac{c\theta}{d} < 1$ and let $\alpha := c\theta$. Then (as $v \to 0$)

$$\begin{aligned} \frac{ds(v)}{dv} &= \exp\left[-\int_\theta^v \tfrac{\alpha}{uh(u)}du\right] \sim v^{-\alpha} \\ s(v) &= \int_0^v \exp\left[-\int_\theta^y \tfrac{\alpha}{wh(w)}dw\right] dy \sim v^{1-\alpha} \end{aligned} \tag{1.14}$$

and

$$m(dy) = I_{(0,\infty)} \frac{1}{4(1-\alpha)^2} \frac{1}{y^{\frac{2\alpha-1}{\alpha-1}}} dy. \tag{1.15}$$

Then by the standard criterion for accessibility ($\int_{0+} ym(dy) < \infty$) ([RW, 87], Theorem 51.2) we obtain $P(\widehat{T}_0 < T) > 0$ for $T > 0$. To obtain (1.5) for $Z(t)$ we use Girsanov to obtain $P(T_0 < T) > 0$ for $T > 0$, the recurrence of the point 1 for (1.1) and the strong Markov property to complete the proof.

In the following lemma we collect a number of key properties related to the set of zeros of $Z(t)$.

Lemma 1.2. *Consider the diffusion (1.1) and assume* $0 < \frac{c\theta}{h(0)} < 1$. *Then define:*

$$I_0 = \{t > 0 | Z_t = 0\}. \tag{1.16}$$

(a) The point 0 *is instantaneously reflecting and the local time at 0 is 0.*
(b) I_0 *has no interior.*
(c) Let λ *denote Lebesgue measure, then*

$$\lambda(I_0) = 0 \quad a.s. \tag{1.17}$$

and

$$P(Z_t = 0) = 0 \text{ for } t > 0. \tag{1.18}$$

(d) The complement of I_0 *is a countable union of disjoint intervals, called the* excursion intervals.
(e) If $Z_0 > 0$*, then the number of zeros in a time interval* $[0, t)$ *is either 0 or infinite.* □

Proof The parts (a)-(d) can be proved using the methods in [RY], XI (1). The results (1.17),(1.18) can also be found in [RW, 87], vol. 2, V.48. The excursion decomposition is discussed in detail in ([RY]). The part (e) follows from the representation of the process in terms of a reflecting Brownian motion given in [RW, 87] in V (48.6).

We now turn to the behavior near a zero. Since Z_t^{-1} occurs as a transition rate in the construction of the dual process it is important to know whether the singularities are integrable or not.

Lemma 1.3. *Consider the diffusion* Z_t *in (1.1) with* $c\theta > 0$*. Let* τ_0 *be a stopping time and* $Z(\tau_0) = 0$*. Then for any* $\varepsilon > 0$*,*

$$\int_{\tau_0}^{\tau_0+\varepsilon} \frac{1}{Z(s)} ds = \infty, \quad a.s. \qquad \square \tag{1.19}$$

Proof We adapt the proof of Lemma 1.6 in chapter 9 [EK]. Define for $0 < t < \varepsilon$,

$$\begin{aligned} \tau(t) &= \inf\left(s > 0 : \int_0^s (Z(\tau_0 + u))^{-1} du = t\right), \quad \text{if } \frac{1}{Z(\tau_0 + \cdot)} \text{ is integrable} \\ \tau(t) &= 0, \quad \text{otherwise} \end{aligned} \tag{1.20}$$

and $\widetilde{Z}(t) = Z(\tau_0 + \tau(t))$. Then we have (see (1.4) in Chapt. 6 of [EK] for a more detailed calculation)

$$\tau(t) = \int_0^t \widetilde{Z}(s) ds, \quad \dot{\tau}(t) = \widetilde{Z}(t), \tag{1.21}$$

which implies the relation

$$E[\tau(t)] = \int_0^t E[\widetilde{Z}(s)] ds = \int_0^t E[Z(\tau(s)) ds] \le c\theta \cdot \int_0^t E[\tau(s)] ds. \tag{1.22}$$

Hence

$$E[\tau(t)] \le E[\tau(0)] e^{c\theta(t)}. \tag{1.23}$$

Since trivially $E[\tau(0)] = 0$, we therefore know $E[\tau(t)] = 0$, $\forall t \ge \tau_0$ which implies that $\tau(t) = 0$ a.s. yielding the result.

When 0 is a recurrent point we will also need to prove that a.s. the same property is true at every zero of the process Z.

Lemma 1.4. *Consider the diffusion* Z_t *in (1.1) with* $0 < c\theta < d$*. Then for any* $\varepsilon > 0$*,*

$$\int_{t_0}^{t_0+\varepsilon} \frac{1}{Z(s)} ds = \infty \quad \forall t_0 \in I_0, \quad a.s. \qquad \square \tag{1.24}$$

Proof First note that since I_0 has no interior, every zero is either the starting time of an excursion or is the limit from the right of a sequence of such times. Since there are only countably many excursions it therefore suffices to prove the result for the starting point of a typical excursion.

The proof is based on the representation of Z as a time change of reflecting Brownian motions (following RW (48.6)).

The scale function is given by

$$u = s(v), \quad v = s^{-1}(u) \tag{1.25}$$

$$\begin{aligned} \frac{ds(v)}{dv} &= \exp\left[-\int_\theta^v \frac{\alpha}{uh(u)}du\right], \\ s(v) &= \int_0^v \exp\left[-\int_\theta^y \frac{\alpha}{wh(w)}dw\right]dy \sim v^{1-\alpha}. \end{aligned} \tag{1.26}$$

We will now construct a version of the process Z via a scale and time change of a reflecting Brownian motion $(U_t)_{t\geq 0}$. Define:

$$V_t = f(U_t) = s^{-1}(U_t). \tag{1.27}$$

Then we have:

$$dV_t = f'(U_t)dU_t + \frac{1}{2}f''(U_t)dt \tag{1.28}$$

Since

$$\frac{\partial s^{-1}(u)}{\partial u} = \left(\frac{\partial v}{\partial u}\right)(u) = \frac{1}{\frac{\partial s}{\partial v}} = \exp\left(\int_\theta^{s^{-1}(u)} \frac{\alpha}{wh(w)}dw\right) \tag{1.29}$$

and

$$\begin{aligned} \frac{\partial^2 s^{-1}(u)}{\partial u^2} &= \frac{\alpha}{s^{-1}(u)h(s^{-1}(u))}\exp\left(\int_\theta^{s^{-1}(u)} \frac{\alpha}{wg(w)}dw\right)\frac{\partial s^{-1}(u)}{\partial u} \\ &= \frac{\alpha}{s^{-1}(u)h(s^{-1}(u))}\exp\left(2\int_\theta^{s^{-1}(u)} \frac{\alpha}{wh(w)}dw\right), \end{aligned} \tag{1.30}$$

we get that

$$\begin{aligned} dV_t &= \frac{1}{\sqrt{V_t h(V_t)}}\exp\left(\int_\theta^{V_t} \frac{\alpha}{xh(w)}dw\right)\sqrt{V_t h(V_t)}dU_t \\ &\quad + \frac{1}{V_t h(V_t)}\exp\left(2\int_\theta^{V_t} \frac{\alpha}{xh(w)}dw\right)\frac{\alpha}{2}dt. \end{aligned} \tag{1.31}$$

We now make a time change as follows: (here ℓ_t^y is Brownian local time)

$$\varphi_t = \int_0^t \frac{1}{V_s h(V_s)}\exp\left(2\int_\theta^{V_s} \frac{\alpha}{xh(w)}dw\right)1(V_s > 0)ds. \tag{1.32}$$

Since near zero the integrand satisfies

$$\sim \frac{1}{4}\frac{1}{(1-\alpha)^2}U_s^{(\frac{1}{1-\alpha}-2)} \tag{1.33}$$

we know that the integrability is the same as the finiteness of

$$\int_0^\infty \frac{1}{4}\frac{1}{(1-\alpha)^2}y^{(\frac{1}{1-\alpha}-2)}\ell_t^y dy. \tag{1.34}$$

This quantity is finite for every t with probability 1 since $\frac{1}{1-\alpha} - 2 > -1$. The inverse is obtained from

$$t = \int_0^{\tau_t} \frac{1}{V_s h(V_s)} \exp\left(2\int_\theta^{V_s} \frac{\alpha}{xh(w)} dw\right) ds. \tag{1.35}$$

Then the process $(V_{\tau_t})_{t\geq 0}$ has the same law as $(Z_t)_{t\geq 0}$ (see RW (48.10)). Consider an excursion of Z_t of length $\geq \varepsilon$. Then it is obtained from a corresponding excursion of U and we can calculate

$$\begin{aligned} \int_0^\varepsilon \frac{1}{Z_s} ds &= \int_0^\varepsilon \frac{1}{V_{\tau_s}} ds \\ &= \int_0^{\tau_\varepsilon} \frac{1}{V_t} \frac{ds}{dt} dt \\ &= \int_0^{\tau_\varepsilon} \frac{1}{V_t^2 h(V_t)} \exp\left(2\int_\theta^{V_t} \frac{\alpha}{xh(w)} dw\right) dt \\ &= \int_0^{\tau_\varepsilon} \frac{1}{(s^{-1}(U_t))^2 h(s^{-1}(U_t))} \exp\left(2\int_\theta^{s^{-1}(U_t)} \frac{\alpha}{xh(w)} dw\right) dt. \end{aligned} \tag{1.36}$$

Near $u = 0$, $s^{-1}(u) \sim u^{\frac{1}{1-2\alpha}}$, $\exp(2\int_\theta^{s^{-1}(u)} \frac{\alpha}{wh(w)} dw) \sim (s^{-1}(u))^{4\alpha} \sim u^{\frac{4\alpha}{1-2\alpha}}$, so that

$$\frac{1}{(s^{-1}(u))^2 h(s^{-1}(u))} \exp\left(2\int_\theta^{s^{-1}(u)} \frac{\alpha}{xh}\right) \sim \frac{1}{u^2}. \tag{1.37}$$

Therefore since an excursion of U near zero is a Bes(3) process(see Rogers, Williams VI, Theorem 55.11),

$$\begin{aligned} \int_0^\varepsilon \frac{1}{Z_s} ds &= \int_0^{\tau_\varepsilon} \frac{1}{(s^{-1}(U_t))^2 h(s^{-1}(U_t))} \exp\left(2\int_\theta^{s^{-1}(U_t)} \frac{\alpha}{xh(w)} dw\right) dt \\ &\sim \int_0^{\tau_\varepsilon} \frac{1}{U_s^2} ds = \infty \end{aligned} \tag{1.38}$$

for the latter see RY, XI Exer. 2.5(4). Hence the proof is complete.

Remark. In the case $g \equiv d$, $\frac{c\theta}{d} = 1$, Z_t is the square of a BES(2) process. In exercise RY 2.6, the same result is proved for BES(2) which arises when $c\theta = d$.

Lemma 1.5. *Let Assume that $\theta(t)$ is continuous, adapted, strictly positive and bounded above by θ^* on $[0,T]$. Let $\tau_{m,n} := \hat{\tau}_{m,n} + \frac{1}{n}$ where $\hat{\tau}_{m,n}$ is the starting point of the mth excursion of length $\geq \frac{1}{n}$. Then*

$$\int_{\hat{\tau}_{m,n}}^{\hat{\tau}_{m,n}+\varepsilon} \frac{1}{Z(s)} ds = \infty \quad \forall m,n \quad a.s. \qquad \square \tag{1.39}$$

Proof Consider the modified processes

$$\theta_{m,n}(t) = 1(\{t \leq \tau_{m,n}\})\theta(t) + 1(\{t > \tau_{m,n}\})\theta^* \tag{1.40}$$

and the solutions $Z_{m,n}$ (with respect to the same Brownian motion) of the equation (1.3) with $\theta(t)$ replaced by $\theta_{m,n}(t)$. Then using the strong comparison theorem we obtain that

$$Z(t) \leq Z_{m,n}(t), \quad a.s. \tag{1.41}$$

Taking limits we get $Z(\hat{\tau}_{m,n} + t)$, $t \geq 0$ is dominated by an excursion of the process with $\theta(t)$ replaced by $1(\{t \leq \hat{\tau}_{m,n}\})\theta(t) + 1(\{t > \hat{\tau}_{m,n}\})\theta^*$. The result then follows from Lemma 1.4.

(ii) Zeros and behavior near zeros for the interacting system (0.37)

The existence and behavior near zeros and *joint* zeros of $\bar{x}_\xi(\cdot)$, $\bar{x}_\eta(\cdot)$ at points $\xi, \eta \in \Omega$ with $a(\xi,\eta) > 0$ has implications for the construction of the dual process we carry out in subsection 1(b). In the present subsection we establish Proposition 0.2 (a) - (g) and in the next one, i.e. 1(b), the part (h) which together provide the necessary results to construct the dual. Throughout this subsection we consider the system (0.37) with stationary ergodic initial conditions with strictly positive intensity. We shall proceed in various steps to carry this out. Here in paragraph (ii) we first verify part (a) and then in one piece parts (b)-(f). The part (g) will be handled after some preparation in (iii) and (iv)

Proof of Proposition 0.2(a)

As an important first step we show that with probability one there exist finite subsets of Ω in which there are no simultaneous zeros in $[0,T]$. As a preliminary indication of why this is true we first prove:

Lemma 1.6. *Let $\{(Z^i_t)_{t\geq 0},\ i \in \mathbb{N}\}$ be an independent collection of diffusions given by (1.1) with $g(x) = dx$. Then for every $c, d, \theta > 0$ fixed there exists an $n_0 \in \mathbb{N}$ such that for $n > n_0$ and i_k all different:*

$$Prob\ (\exists t \in (0,\infty) : Z^{i_1}_t = Z^{i_2}_t = \ldots Z^{i_n}_t = 0) = 0 \qquad \square \tag{1.42}$$

Proof Consider the stochastic differential equation for the sum of n-components, denoted $S_n(t)$, which is in law equivalent to

$$dS_n(t) = c(n\theta - S_n(t))dt + \sqrt{2d \cdot S_n(t)}dw(t). \tag{1.43}$$

Now apply (1.4) to get the result for

$$n_0 = [d/(2c\theta)]. \tag{1.44}$$

In the interactive case we get the following weaker result which yields Proposition 0.2(a).

Lemma 1.7. *Let $\{(\bar{x}_\xi(t))_{t\geq 0},\ \xi \in \Omega\}$ be the collection of diffusions given by (0.37) satisfying the hypothesis of Proposition 0.2 and Ω_n an increasing sequence of finite subsets of Ω. Then for a.e. $\bar{X}$ there exists $n_0(T, \bar{X}(0))$ such that for $n \geq n_0(T, \bar{X}(0))$,*

$$\sum_{\xi\in\Omega_n} \bar{x}_\xi(t) \neq 0 \quad \forall t \in [0,T]. \qquad \square \tag{1.45}$$

Remark Lemma 1.6 proves non-percolation, in fact, the non-existence of an infinite number of simultaneous zeros in the mean-field case. In the interactive case the existence or non-existence of an infinite connected set of zeros is open.

Proof We consider the process

$$S_n(t) = \sum_{i=1}^n \bar{x}_{\xi_i}(t). \tag{1.46}$$

Then we have with $c = \sum_\xi a(0,\xi)$ that:

$$dS_n(t) = c\left(\sum_{i=1}^n \theta_i(t) - S_n(t)\right)dt + \sum_{i=1}^n \sqrt{2g(\bar{x}_{\xi_i}(t))}dw_{\xi_i}(t). \tag{1.47}$$

¿From (1.47) we have

$$S_n(t) \geq S_n(0) - c\int_0^t S_n(t)ds + \sum_{i=1}^n \int_0^t \sqrt{g(\bar{x}_{\xi_i}(s))}dw_{\xi_i}(s). \tag{1.48}$$

Therefore,

$$\frac{S_n(t)}{n} \geq \frac{S_n(0)}{n}e^{-ct} + e^{-ct}\int_0^t e^{cs}\frac{1}{n}\sum_{i=1}^n \sqrt{g(\bar{x}_{\xi_i}(s))}dw_{\xi_i}(s). \tag{1.49}$$

Note that the increasing process of the stochastic integral above is (due to the sublinearity of h) close to 0 bounded above in expectation by $\frac{D}{n^2}E[\int_0^t e^{2c}S_n^2(s)ds]$ for some finite D independent of n, which implies convergence to 0 in mean. Since the initial state is translation invariant, ergodic with positive intensity we conclude:

$$\lim_{n\to\infty} \frac{S_n(0)}{n} > 0. \tag{1.50}$$

From this, the above bound on the increasing process and the martingale inequality applied to the stochastic integral we get

$$\lim_{n\to\infty} P(\sup_{t\in[0,T]} |\int_0^t e^s \frac{1}{n}\sum_{i=1}^n \sqrt{g(\bar{x}_{\xi_i}(s))}dw_{\xi_i}(s)| \geq \frac{S_n(0)}{n}) = 0. \tag{1.51}$$

The result of Proposition 0.2 part(a)then follows by a standard Borel-Cantelli argument.

Proof of Proposition 0.2, (b)-(f).

Now we want to use the result in Lemma 1.2 to obtain the corresponding assertions for the interacting system. Consider a finite subset Ω_n in which there are no simultaneous zeros. Then we can find a finite number of closed intervals of positive length, $\{[t_i, t_{i+1}] : i = 1, \dots, m\}$ covering $[0, T]$ and elements $\{\xi_i,\ i = 1, \dots, m\}$ such that $\bar{x}_{\xi_i}(t) > 0, \forall t \in [t_i, t_{i+1}]$. Now consider a typical time interval $[t_i, t_{i+1}]$. We can obtain (0.39), (0.40), (0.41) for $\{\xi : \rho(\xi, \xi_i) \leq 1\}$ (recall (0.7)) using Lemma 1.2. Then one obtains the analogous results for $\{\xi : \rho(\xi, \xi_i) \leq k\},\ k = 2, \dots, |\Omega_n|$ by induction. Hence we have shown Proposition 0.2 parts (b)-(f).

(iii) Behavior at a connected joint zero set - analysis of a finite system caricature.

Now we still have to prepare for the proof of proposition 0.2 part (g). In the presence of simultaneous 0's of the total mass process at different sites, the ratio of total masses at two such sites enter in the construction of the dual process. Consequently we need to determine the integrability of these ratios as the *joint* zero is approached. In order to introduce the essential ideas we first consider a simplified finite system which serves as a caricature of the general case. Then in the next subsection we indicate how the same arguments can be applied to the system (0.37).

Consider the finite system of stochastic differential equations:

$$\begin{array}{lll} dz_0(t) & \equiv \theta > 0 & \\ dz_1(t) & = (c_1 z_0(t) - b_1 z_1(t))dt & + \sqrt{2h(z_1(t))z_1(t)}dw_1(t) \\ dz_2(t) & = (c_2 z_1(t) - b_2 z_2(t))dt & + \sqrt{2h(z_2(t))z_2(t)}dw_2(t), \\ dz_3(t) & = (c_3 z_2(t) - b_3 z_3(t))dt & + \sqrt{2h(z_3(t))z_3(t)}dw_3(t), \\ \dots & \dots & \dots \\ dz_n(t) & = (c_n z_{n-1}(t) - b_n z_n(t))dt & + \sqrt{2h(z_n(t))z_n(t)}dw_n(t), \end{array} \tag{1.52}$$

where $\theta > 0$ and $c_i, b_i > 0, \quad i = 1, \dots, n$.

As above this system can exhibit joint zeros, that is, times t_0 at which $z_1(t_0) = z_2(t_0) = \dots = z_n(t_0) = 0$. In the following two lemmas we identify the behavior in the neighborhood of such a joint zero, first for the case of a fixed time or a stopping time and then for the case of a zero at which an excursion away from zero begins.

Consider first a time point t_0 (fixed time or a stopping time) such that

$$z_1(t_0) = z_2(t_0) = \dots = z_n(t_0) = 0. \tag{1.53}$$

Then we claim that on this event the following holds:

Lemma 1.8. *Under the assumptions (1.52) and (1.53): for any $s_0 > 0$, and $1 \leq j \leq n$,*

$$\int_{t_0}^{t_0+s_0} \frac{z_{j-1}(s)}{z_j(s)} ds = +\infty \ a.s. \qquad \square \tag{1.54}$$

Proof The result for $j = 1$ follows from Lemma 1.3. For $j = 2, \dots, n$ our strategy of proof will be to show that arbitrarily close to t_0 there are points τ_m such that $z_j(\tau_m) = 0$ and $z_{j-1}(\tau_m) > 0$, so that for every $\varepsilon > 0$ the integral from τ_m to $\tau_m + \varepsilon$ is infinite by Lemmas 1.3,1.4.

We begin with some preliminaries. Using a simple time change there is no loss of generality in assuming that $t_0 = 0$ and $h(0) = \frac{1}{2}$ (recall zeros can only appear if $h(0) > 0$). Using (1.18) and the trajectorial continuity of z_{j-1} we can prove inductively that

$$P(z_j(t) = 0) = 0 \text{ for all } t > 0. \tag{1.55}$$

Consider next the strong solution of the stochastic differential equation

$$dy_j(t) = \frac{2h(y_j(t))}{4}dt + \sqrt{2h(y_j(t))y_j(t)}dw_j(t), \quad y_j(0) = 0, \tag{1.56}$$

where we use the same Brownian motions $(w_j(t))_{t \geq 0}$, as in (1.52). Next let

$$v_j(t) := 2\sqrt{y_j(t)}. \tag{1.57}$$

Then by Ito's formula

$$dv_j(t) = \sqrt{2h(\frac{v_j^2}{4})}dw_j(t). \tag{1.58}$$

Now let

$$\tau := \inf\{t > 0 : c_j z_{j-1}(t) - b_j z_j(t) = \frac{1}{8}\} \wedge \inf\{t > 0 : h(y_j(t)) \leq \frac{1}{4}\}. \tag{1.59}$$

By the coupling of the z and the y-system using the same Brownian motions $z_j(t) \leq y_j(t)$ on $\{t \leq \tau\}$. For $m \geq 1$, let

$$\tau_m = \inf\{t > \frac{1}{m} : y_j(t) = 0\}. \tag{1.60}$$

Then $z_j(\tau_m) = y_j(\tau_m) = 0$ on $\{\tau_m \leq \tau\}$. But, τ_m is independent of $w_1(\cdot), \dots, w_{j-1}(\cdot)$, so $z_{j-1}(\tau_m) > 0$ since $P(z_{j-1}(t) = 0) = 0$ for any $t > 0$ by (1.55). Using Lemma 1.3, we then have for $\varepsilon > 0$

$$\int_{\tau_m}^{\tau_m+\varepsilon} \frac{z_{j-1}(t)}{z_j(t)}dt = \infty \quad \text{a.s. on } \{\tau_m \leq \tau\}. \tag{1.61}$$

Since $v_j(t)$ is a time change of Brownian motion, it takes value zero infinitely often on $[0, \delta]$ for every $\delta > 0$ and since $\tau > 0$ a.s., we have

$$P(\cup_{m=1}^{\infty}\{\tau_m \leq \tau, \tau_m < \delta\}) = 1. \tag{1.62}$$

This together with (1.61) yields

$$\int_0^{\delta} \frac{z_{j-1}(t)}{z_j(t)}dt = \infty \quad \text{a.s.} \tag{1.63}$$

which completes the proof.

The above result shows that for fixed times t_0 no excursions of total mass away from zero occur. However they do occur along the path. We now consider a time point where an excursion of $\sum_{i=1}^{n} z_i(t)$ away from 0 starts and obtain a similar result.

Lemma 1.9. *Consider the system of SDE's given in (1.52) and let* τ_0 *be the beginning of an excursion of* $\sum_{i=1}^{n} z_i(t)$ *away from 0. Then for* $s_0 > 0$,

$$\int_{\tau_0}^{\tau_0+s_0} \frac{z_{j-1}(s)}{z_j(s)} ds = +\infty \ a.s. \quad \square \tag{1.64}$$

Proof We first observe that the result for $j = 1$ follows immediately from Lemma 1.4 since $z_0(s) \equiv \theta$ and $z_1(s)$ has a singularity at $s = \tau_0$.

To prove the result for $j = 2, \dots, n$ we will show that each $z_j, j = 2, \dots, n$ has countably many zeros in $(\tau_0, \tau_0 + s_0) \cap \{s : z_{j-1}(s) > 0\}$ and then again the result follows from Lemma 1.4.

Define:

$$R(t) = \sum_{i=1}^{n} z_i(t) \tag{1.65}$$

and consider the process:

$$(z_i^*(t))_{i=0,1,\dots,n} \quad \text{with} \quad z_i^*(t) = z_i(t)/R(t) \quad i = 1, \dots, n. \tag{1.66}$$

Let $\mathcal{F}_t^n$ be the σ-algebra generated by $\{w_i(s) : 0 < s \le t, \ i = 1, \dots, n\}$. Then we can represent $R(t)$ by the following Ito equation:

$$dR(t) = [(c_1\theta - b_1 z_1(t)) + \sum_{i=2}^{n} (c_i z_{i-1}(t) - b_i z_i(t))]dt + \sqrt{\sum_i h(z_i(t)) z_i(t)} dw^*(t) \tag{1.67}$$

where $(w^*(t))_{t\ge0}$ is a $(\mathcal{F}_t^n)_{t\ge0}$ adapted Brownian motion. Since we assume that $R(\tau_0) = 0$, we conclude by Lemma 1.3 that

$$\int_{\tau_0}^{\tau_0+s_0} (R(t))^{-1} dt = +\infty. \tag{1.68}$$

Next we derive the evolution equation for the z^*-process. By a standard calculation using Ito's lemma we obtain that $(z_i^*(t))_{t\ge0}, \ i \in \{2, \dots, n\}$ is a semi martingale satisfying the following system of Ito equations (we suppress now t in $z_i(t)$ and $R(t)$):

$$\begin{aligned} dz_j^* &= \Big[(1 - z_j^*)(c_j z_{j-1}^* - b_j z_j^*) - z_j^*(\textstyle\sum_{i\ne j,1}(c_i z_{i-1}^* - b_i z_i^*)) - z_j^*(\frac{c_1\theta}{R} - b_1 z_1^*)\Big] dt \\ &\quad -2z_j^* \frac{[\sum_{i\ne j} z_i^*(h(z_j) - h(z_i))]}{R} dt \\ &\quad + \frac{(1-z_j^*)\sqrt{h_j z_j^*} dw_j - z_j^*(\sum_{i\ne j} \sqrt{h_i z_i^*} dw_i)}{R^{1/2}}, \qquad j = 1, \dots, n. \end{aligned} \tag{1.69}$$

Consider first the case where $h(x) \equiv h > 0$ in order to understand the basic idea. In this situation we can verify as in Section 2, Lemma 2.6, that *conditioned on* $(R(t))_{t\ge0}$, the normalized process $(z_1^*(t), \dots, z_n^*(t))$ is a time-inhomogeneous *Fisher-Wright diffusion with n-types* (i.e. Fleming-Viot process on the type set $\{1, 2, \dots, n\}$), where the resampling rate is $(R(t))^{-1}h$ and mutation operator given by the drift terms in (1.69). Similarly, conditioned on $(R(t))_{t\ge0}$, for $j > 1$, the component $z_j^*(t)$ is a time-inhomogeneous Wright-Fisher diffusion with diffusion coefficient $(R(t))^{-1}h$. Namely it satisfies the following Ito equation:
There exists a process $(u(t))_{t\ge0}$ which is $(\mathcal{F}_t^n)_{t\ge0}$-adapted and bounded above on $(0, s_0)$, such that

$$dz_j^* = \left[c_j z_{j-1}^* + u z_j^*\right] dt + \frac{\sqrt{h(1 - z_j^*) z_j^*} dw_j^{**}}{R^{1/2}}.$$

Here $(w_j^{**}(t))_{t\ge0}$ is a Brownian motion, which is $(\mathcal{F}^{(n)})_{t\ge0}$ adapted. Moreover the term $z_{j-1}(t) = R(t) z_{j-1}^*(t)$ is adapted, positive for a.e. t and independent of $\{w_j(t), t \ge 0\}$. Then applying a time change and argument as in the proof of Lemma 1.8 it follows that z_j^* has a zero in the open set $(0,s_0)\cap\{t: z_{j-1}^*>0\}$. Then (1.64) follows from Lemma 1.4.

The general case can be obtained by using the Lipschitz property of h at 0 to show that the drift term is of the same form and the diffusion rate in $(0,\varepsilon)$ is given by replacing h in the second term in (1.70) above by $v(t)$ where $(v(t))_{t\geq 0}$ is $(\mathcal{F}_t^*)_{t\geq 0}$-adapted and bounded below by $\delta>0$ on $(0,s_0)$.

(iv) Extension to Interacting Systems

Now we have the tools to treat the system of interest and prove another part of Proposition 0.2.

Proof of Proposition 0.2(g)

Let $\{(\bar{x}_\xi(t))_{t\geq 0},\ \xi\in\Omega\}$ be the collection of diffusions given by (0.37) satisfying the hypothesis of Proposition 0.2.

Lemma 1.10. *Consider the system (0.37) with* $c\equiv 0$ *and suppose that* $\mathcal{N}$ *is a subset of a finite communicating (w.r.t.* $a(\cdot,\cdot)$*) set of sites* $\mathcal{N}_+$ *such that there are no common zeros in* $\mathcal{N}_+$ *in* $[0,T]$*. Assume that* $\sum_{\eta\in\mathcal{N}}\bar{x}_\eta(\tau_0)=0$ *and that* $\tau_0\in[0,T]$ *is either a stopping time or the starting time of an excursion away from zero of* $\bar{x}_\xi(t)$ *with* $\xi\in\mathcal{N}$*. Then there exists* $\eta\in\mathcal{N}_+$ *with* $a(\xi,\eta)>0$ *such that*

$$\int_{\tau_0}^{\tau_0+s_0}\frac{\bar{x}_\eta(s)}{\bar{x}_\xi(s)}ds=+\infty\ a.s. \qquad \square \tag{1.70}$$

Proof The proof proceeds by arguments parallel to those given in Lemmas 1.8 and 1.9 and hence we sketch the proof only. Namely we indicate the modifications required to handle the general system (0.37). First note that since we need only consider the behavior close to τ_0 we can freeze the values of $\bar{x}_\zeta(t),\ \tau_0-\varepsilon\leq t\leq\tau_0+\varepsilon,\quad \zeta\notin\mathcal{N}$ and consider the finite system restricted to $\mathcal{N}$ with fixed positive immigration given by these frozen values.

Now recall that we have assumed that all sites communicate in the sense that between two points ξ,η there is a sequence $\xi_1,\dots,\xi_k$ with $a(\xi,\xi_1)>0,\dots,a(\xi_k,\eta)>0$. Hence $\mathcal{N}=\cup_{j=1}^n\mathcal{N}_j$ for some finite n where $\zeta\in\mathcal{N}_j$ if $\sum_{k=1}^j a^k(\eta,\zeta)>0$ for some $\eta\in\mathcal{N}^c$. The difference between $\mathcal{N}_{m-1}$ and $\mathcal{N}_m$ is denoted by $\partial\mathcal{N}_m$.

In the case where $a(\xi,\eta)>0$ for a infinite collection of η or if ξ is an isolated zero, then $\mathcal{N}=\mathcal{N}_1$ and nothing needs to be shown since then there will be an η with $a(\xi,\eta)>0$ and $x_\eta(t_0)>0$ and the claim follows by Lemma 1.3. Hence in the sequel we assume that ξ is a member of a finite connected set, $\mathcal{N}=\cup_{j=1}^m\partial\mathcal{N}_j$, with $m>1$ and that $a(\xi,\cdot)$ has *finite support*.

There are again two cases to consider, namely, the case in which τ_0 is a fixed time (or stopping time) and the case in which τ_0 is the initial point in an excursion away from zero of $\sum_{\zeta\in\mathcal{N}}\bar{x}_\zeta(t)$. The proofs of Lemmas 1.8 and 1.9 can be adapted in two steps as follows.

We first modify the systems there to allow migration from i to $i-1$ (in addition to $i+1$). This does not effect the proof of the existence of zeros. Moreover, to verify the existence of zeros in $\{t: z_{j-1}^*(t)>0\}$ we use a comparison argument where we compare with the system with the migration $j+1$ to j removed so that the independence of z_{j+1}^* from w_j holds. The second step involves replacing the component z_i^* by the collection

$$\{\bar{x}_\zeta^*(t):\zeta\in\partial\mathcal{N}_i\},\quad where\quad \bar{x}_\zeta^*(t):=\frac{\bar{x}_\zeta(t)}{\sum_{\eta\in\mathcal{N}}\bar{x}_\eta(t)}. \tag{1.71}$$

If we first assume that there is no immigration from $\partial\mathcal{N}_{i+1}$ to $\partial\mathcal{N}_i$, then it can then be verified that $\bar{x}_\zeta^*(t)$ with $\zeta\in\partial\mathcal{N}_j$ satisfies an equation of the same form as (1.70) (but with bounded immigration from other sites in $\partial\mathcal{N}_j$). We can then deduce the existence of a zero in $(0,s_0)\cap\{t:\sum_{\eta\in\partial\mathcal{N}_{j-1},a(\eta,\zeta)>0}\bar{x}_\eta^*(t)>0\}$ as in the proof of Lemma 1.9. The same result allowing migration from $\partial\mathcal{N}_{i+1}$ to $\partial\mathcal{N}_i$ is again obtained from a comparison argument. Finally, the result follows by another application of Lemma 1.3.

Now we can return to the proof of Proposition 0.2. Given $\xi\in\Omega$ and $\tau_0\in[0,T]$ with $\bar{x}_\xi(\tau_0)=0$ (with τ_0 a stopping time or starting point of an excursion of $\bar{x}_\xi(t)$ away from zero), part (a) of the Proposition implies that we can choose a finite (random) set $\mathcal{N}_+$, with $\xi\in\mathcal{N}_+$, so that there are no common zeros in $\mathcal{N}_+$ in $[0,T]$. Part (g) of Proposition 0.2 then follows from the previous lemma.

(b) Random walk and coalescing random walk in a singular random environment. In this subsection we construct a system of coalescing random walks on Ω described by a Markov jump chain in which the jump rates are possibly unbounded. We carry out this construction in two steps. We first construct a Markov jump process on Ω with singular rates and then using this, we construct the system of coalescing random walks with singular coalescence rates. In a third step we apply this to our concrete situation.

The rates will be built from collections

$$\{(m_\xi(t))_{t\geq 0},\ (d_\xi(t))_{t\geq 0},\ \xi\in\Omega\} \tag{1.72}$$

which will in fact now for our application also depend on a parameter T which we suppress whenever possible without confusion. We begin by fixing $0<T<\infty$ and specifying in $(\alpha)-(\beta)$ the parameters and those properties relevant for our construction:

(α) There is a family of jump rates $\tilde{a}^{T,t}(\xi,\eta)$ on $\Omega\times\Omega$ for every time $0\leq t\leq T$, which are of the form:

$$\tilde{a}^{T,t}(\xi,\eta)=m_\xi(t)a(\xi,\eta)m_\eta^{-1}(t),\quad t\in[0,T],\quad \xi\neq\eta, \tag{1.73}$$

where $a(\xi,\eta)$ is a random walk transition rate on $\Omega\times\Omega$. The collection $\{(m_\xi(t))_{t\geq 0}\}_{\xi\in\Omega}$ has the following property (here λ denotes the Lebesgue measure):

$$\begin{aligned}
&-t\to m_\xi(t)\quad \text{is a function } [0,\infty)\to(0,\infty],\ \text{which is continuous on } \bar{\mathbb{R}}^+,\\
&-\text{the singularities are not integrable from either side},\\
&-\lambda\{t\leq T: m_\xi(t)=\infty\}=0\quad \forall T>0,\ \xi\in\Omega\\
&-\sharp\{t|t\leq T,\ m_\xi(t)=\infty\quad \forall\xi\in I\}=0\quad \text{for every } I \text{ with } |I| \text{ countable },\ T<\infty.\\
&-\text{There is for every }\ \xi \text{ and }\ t_0 \text{ with }\ m_\xi(t_0)=\infty \text{ an } \eta \text{ with}\\
&\qquad a(\xi,\eta)>0 \text{ such that }\ m_\xi(t_0)/m_\eta(t_0)\ \text{ is not integrable at }\ t_0.
\end{aligned} \tag{1.74}$$

(β) The coalescence rates of the Markov chain are determined by a collection $\{(d_\xi(t))_{t\geq 0}\}_{\xi\in\Omega}$ of continuous functions from $[0,\infty)$ into $[0,\infty]$, which satisfy also the condition (1.74) except the last one which is replaced by requiring that for every singularity t_0 at the point ξ:

$$d_\xi(t)/m_\xi(t)\quad \text{converges as } t\to t_0. \tag{1.75}$$

These are all the ingredients we need for our construction in the next two subsections.

(i) Random walk with singular rates

We want to define for every T a Markov process $(\xi(t))_{t\geq 0}$ with rates $\tilde{a}^{T,t}(\xi,\eta)$ given by

$$\tilde{a}^{T,t}(\xi,\eta)=m_\xi(t)a(\xi,\eta)(m_\eta(t))^{-1},\quad t\in[0,T], \tag{1.76}$$

and with starting point $\xi(0)=\zeta$ satisfying $m_\zeta(0)<\infty$.

We first note that if $\sup\{m_\xi(t):\xi\in\Omega,t\in[0,T]\}<\infty$, then $\xi(t)$ is a càdlàg process on Ω given by the solution of the well-posed martingale problem, requiring that $(M_f(t))_{t\in[0,T]}$ is a martingale for every bounded function f on Ω, with the definition:

$$M_f(t)=f(\xi(t))-f(\xi(0))-\int_0^t\sum_\eta \tilde{a}^{T,s}(\xi(s),\eta)[f(\eta)-f(\xi(s))]ds. \tag{1.77}$$

In order to deal with singularities for the functions $m_\xi(\cdot)$ and $d_\xi(\cdot)$ we add a cemetery state Δ and consider the state space $\Omega\cup\{\Delta\}$ for the Markov chain inducing the locations. If we start outside a singularity we get by (1.77) an unique process up to random times given by hitting those points which cause m_ξ to fall outside $[0,M]$, for M chosen with $M>0$. Moreover, conditions (1.74) imply that the random walk cannot hit a singularity after only finitely many jumps. Given $\{m_\xi,\xi\in\Omega\}$ and ξ a regular point at time 0 let $A_t^{\bar{X}}(\xi,\eta)$ be the minimal transition function of the Markov chain, $\xi(t)$, with transition rates from ξ to η at time t given by $\tilde{a}^{T,t}$.

Then in order to define a transition *probability*, for the Markov chain for all times we set $\xi(t)=\Delta$ if $\xi(s)$ has made countably many jumps in $[0,t]$. Now we have a well-defined process on $\Omega\cup\{\Delta\},t\in[0,T]$ if we start in *regular* locations. We set for the case $m_\xi(0) = +\infty$

$$\tilde{a}^{T,0}(\xi, \Delta) = \infty, \tag{1.78}$$

that is, ξ is an instantaneous state, if $m_\xi(0) = +\infty$.

Remark An alternative approach to the case in which $\xi(0) = \eta,\ \ m_\eta(0) = \infty$ is by defining an entrance law and allowing the Markov chain to make an instantaneous jump according to an entrance law and requiring that the law of $\xi(t)$ be a continuous function of t.

(ii) The coalescing random walk process

We next introduce an associated *system of time-inhomogeneous coalescing random walks* with *delay* that will serve as the dual process in the next subsection. We shall define the process for rates with singularities using the collection of $\{(m_\xi(t))_{t\geq 0},\ (d_\xi(t))_{t\geq 0},\ \xi \in \Omega\}$ satisfying (1.73) to (1.74). This means that we can use the fact that we have in the previous subsection constructed the underlying random walk.

Let

$$\Pi_n = \text{set of partitions of } \{1, \dots, n\}. \tag{1.79}$$

For $\pi \in \Pi_n$, let $|\pi|$ denote the number of (nonempty) partition elements and represent π as a mapping from $\{1, \dots, n\}$ onto $\{1, \dots, |\pi|\}$, in other words, π^{-1} gives the indices of the particles within the partition elements. Each partition element is a subset of $\{1, \dots, n\}$. We can order the partition elements according to the rank of the smallest number it contains.

In the sequel we shall write an element $\eta \in (\Omega \cup \Delta)^n$ as $(\eta_1, \dots, \eta_n)$. We shall define a Markov process:

$$(\zeta_t)_{t\in\mathbb{R}^+} = ((\eta_1(t), \dots, \eta_{|\pi_t|}(t));\ \ \pi_t) \tag{1.80}$$

with state space

$$\mathfrak{Z}_n := \{(\eta_1, \dots, \eta_{|\pi|}; \pi) | \eta_j \in (\Omega \cup \{\Delta\}),\ j = 1, \dots, |\pi|,\ \pi \in \Pi_n\}. \tag{1.81}$$

Here $\eta_i(t)$ denotes the *position in space* of the i-th partition element at time t and π_t describes the particles belonging at time t to the different partition elements. We shall now describe the evolution of this *time-inhomogeneous* Markov process on $\mathfrak{Z}_n$, with possible *singularities* in the rates.

Definition 1.1. The law of the process $(\zeta_t)_{t\in[0,T]}$ is determined by prescribing the mechanism of the evolution as follows:

- (i) Each partition element, independent of the others, performs a continuous time Markov jump process on $\Omega \cup \{\Delta\}$, with jump rates $\tilde{a}^{T,t}(\cdot, \cdot)$ at time t and Δ is defined as trap.
- (ii) All pairs of partition elements, whose location in $(\Omega \cup \{\Delta\})$ coincides at time t and is equal to ξ, coalesces into one partition element at the exponential rate $d_\xi(t)$ and we set $d_\Delta(t) \equiv +\infty$, here the value $+\infty$ means instantaneous transition. $\square$

This process is well-defined, namely:

Lemma 1.11. *The transition rates and the condition given above specify a unique Markov process,* $(\zeta_t)_{t\in[0,T]}$*, on* $\mathfrak{Z}_n$. $\square$

Proof It follows from subsection 1(b)(i) that the motion of the partition elements following independent copies of the random walk with rates $\tilde{a}^{T,t}$ is well-defined.

It thus remains to consider the coalescence component of the dynamics. Note that singularities of $\{d_\xi(t),\ \xi \in \Omega\}$ do not pose any problem, since the coalescent process starts with finitely many particles and every transition reduces the total number of particles by one and hence no explosion can occur. Moreover more than one partition element can never be present at a singularity at ξ at time t_0, i.e. $d_\xi(t_0) = \infty$, because coalescence of all partition elements at a site ξ will occur before t_0 since $\int_0^{t_0} d_\xi(s)ds = \infty$ by (1.74). Moreover, joint singularities of $\tilde{a}^T, t$ and the $d_\xi(t)$ are no problem by (1.75).

(iii) Application to rates given by the total mass process.
Now we complete the proof of Proposition 0.2(h) by applying the previous results to rates given by the total mass process.

In subsection 1(a) we have verified that for, almost every realization of $\bar{X}$, (α) and (β) are satisfied when for fixed $T > 0$, we set:

$$m_\xi(t) = (\bar{x}_\xi(T-t))^{-1}, \quad 0 \le t \le T \tag{1.82}$$

and

$$d_\xi(t) = \frac{h(\bar{x}_\xi(T-t))}{\bar{x}_\xi(T-t)}. \tag{1.83}$$

and hence all results of subsection (i) and (ii) can be applied in particular in our case.

What more can we say about our coalescing random walk with singular rates if the latter are given by the total mass process? We write now $A_t^{\bar{X}}$ for the minimal transition function. The purpose of the next Lemma is therefore to show that for almost every realization of $\bar{X}$, with probability one, our time inhomogeneous Markov chain $(\xi(t))_{t\ge 0}$ does *not* reach Δ in finite time and the Markov chain $(t, \xi_t)_{t\ge 0}$ is well-defined on the state space given by the subset of regular points in $[0,\infty) \times \Omega$. We denote by $N_{T,\zeta}^{\bar{X}}$ the number of jumps occurring in time $[0,T]$ if $\xi(0) = \zeta$.

Lemma 1.12. *Assume that the rates arise via (1.82) and (1.83). Then:*
(a) For almost every realization of $\bar{X}$ and $\zeta \in \Omega$ there is a càdlàg version of $(\xi_t)_{t\ge[0,T]}$ starting at $\zeta \ne \Delta$ such that

$$\sum_{\eta \ne \Delta} A_T^{\bar{X}}(\zeta,\eta) = 1, \quad \forall \zeta \ne \Delta \ a.s. \tag{1.84}$$

and

$$N_{T,\zeta}^{\bar{X}} < \infty \quad a.s. \tag{1.85}$$

(b) If $\zeta_0 = ((\eta_1(0), \dots, \eta_{|\pi_0|}(0); \pi_0)$ and $\eta_i(0) \ne \Delta,\ i = 1, \dots, |\pi_0|$, then ζ_t has locations which do not lie in the cemetery $\triangle$ for all $t \le T$. $\square$

Proof of Lemma 1.12 First note that (1.85) follows immediately from (1.84). The proof of (1.84) will follow from a consistency consideration. Let $a_t(\cdot,\cdot)$ denote the transition probabilities of the continuous time random walk with (time and space homogeneous) rate of a jump from ξ to η, being $a(\xi,\eta)$. The proof of the result is based on the relations which are proved in section 2 between the random walk kernels a_T and $A_T^{\bar{X}}$ and any solution to the martingale problem for X. Using this we obtain the following formula:

Lemma 1.13. *Denote by $E_{\bar{X}}$ the integration with respect to the law of the total mass process. Then for any measurable subset, A of $[0,1]$,*

$$E_{\bar{X}}\left[\sum_{\eta\ne\Delta} A_T^{\bar{X}}(\zeta,\eta)\hat{x}_\eta(0)(A)\bar{x}_\zeta(T)\right] = \sum_{\eta\in\Omega} a_T(\zeta,\eta)x_\eta(0)(A). \quad \square \tag{1.86}$$

Before proving Lemma 1.13 we return to the proof of Lemma 1.12. Since $A_T^{\bar{X}}$ is a sub probability kernel on Ω and $E\bar{x}_\zeta(T) = \sum_\eta a_T(\zeta,\eta)\bar{x}_\eta(0)$ the following inequality:

$$E_{\bar{X}}\big[\sum_\eta A_T^{\bar{X}}(\zeta,\eta)\hat{x}_\eta(0)([0,1])\bar{x}_\zeta(T)\big] \le \sum_\eta a_T(\zeta,\eta)\bar{x}_\eta(0), \tag{1.87}$$

with equality if and only if

$$\sum_{\eta:\bar{x}_\eta(0)\neq 0} A_T^{\bar{X}}(\zeta,\eta)=1, \quad \bar{X} \text{ a.s..} \tag{1.88}$$

Therefore Lemma 1.13 immediately implies that (1.84) is satisfied for almost every realization of $\bar{X}$ and completes the proof of Lemma 1.12 part (a). However by construction of the coalescing random walk we have now also part (b).

Proof of Lemma 1.13 Introduce for the $\widehat{X}$ process a cemetery type $\tilde{\Delta}$ for each component where $\tilde{\Delta} \notin [0,1]$. We also set $\hat{x}_\xi = \delta_{\tilde{\Delta}}$ if $\xi = \Delta$. We calculate on the one hand the expectations of $x_\zeta(T)(A) = \bar{x}_\zeta(T)\widehat{x}_\zeta(T)(A)$ for every $A \in \mathcal{B}([0,1])$ by conditioning on $(\bar{X}(t))_{t\geq 0}$ and then using the first moment calculation for $\widehat{X}(T)$, which gives the formula

$$E[\hat{x}_\zeta(T)(A)|\bar{X}(t)_{t\geq 0}] = \sum_{\eta\neq\Delta} A_T^{\bar{X}}(\zeta,\eta)(\hat{x}_\eta(0)(A)). \tag{1.89}$$

The first moment calculation used here follows from the martingale problem characterization of $\widehat{X}$, conditioned on the complete total mass process, proved in chapter 2(b), step 2. For any solution of this martingale problem the expectation of the relative weights process satisfies a system of differential equations (with singularities). The minimal solution to that system is the r.h.s. of (1.89). Next multiply by $\bar{x}_\zeta(T)$ and then integrate over $(\bar{X}(t))_{t\geq 0}$ to get the l.h.s. of (1.86). On the other hand a standard first moment calculation for any solution to the martingale problem for X yields

$$E[x_\zeta(T)(A)] = \sum_{\eta\in\Omega} a_T(\zeta,\eta)x_\eta(0)(A) \tag{1.90}$$

which is the r.h.s. of (1.86).

We can summarize this paragraph by saying that the random walk in singular medium given by $(\tilde{a}_s^T)_{s\leq T}$, which arises from the total mass process, is a well-defined Markov process with càdlàg paths in the set of regular points, that it the process does not hit singularities. This completes the proof of Proposition 0.2(h).

(c) Dual processes for the interacting Fleming-Viot system with singular rates. For processes with interacting and measure valued components a useful duality relation can often be established. This is the case with a system of interacting Fleming-Viot processes with singular rates as is (as we shall prove later on in section 2) our process $(\widehat{X}(t))_{t\geq 0}$ (conditioned on the $\bar{X}$-process). Some care is needed to deal with the *time inhomogeneity* of the rates of the $\widehat{X}$ process as well as with the *singularities* in the rates. In this section we shall obtain a time-inhomogeneous dual process with singular rates and use it to establish uniqueness in law for the $\widehat{X}$ process.

The formulation of a duality relation requires two ingredients: a dual process, and a class of functions whose expectations determine the law of the process. In subsection 1(b) we already introduced the candidate for the dual process. We next introduce the class of functions in (i) and then in (ii) establish the duality relation between the system $\widehat{X}(t)$ of *time-inhomogeneous* interacting Fleming-Viot processes and the coalescing random walk (with possibly singular rates) $(\zeta(t))_{t\geq 0}$ constructed in subsection 1(b)(ii).

(i) The function class for duality

We consider a solution $(\widehat{X}(t))_{t\geq 0}$ to the martingale problem given in (0.51). We begin by enlarging the type and state space of the process $\widehat{X}(t))_{t\geq 0}$ by adding a type $\widetilde{\Delta}$ and a site Δ. Set

$$E := [0,1] \cup \{\tilde{\Delta}\} \text{ and } \widehat{\Omega} := \Omega \cup \{\Delta\} \tag{1.91}$$

where $\{\widetilde{\Delta}\}$ is taken as an isolated point. Then the state space for $\widehat{X}$ is $(\mathcal{P}(E))^{\widehat{\Omega}}$.

Furthermore we put for all $t \geq 0$

$$x_{\Delta}(t) \equiv \delta_{\tilde{\Delta}}. \tag{1.92}$$

We are interested in characterizing the law of $\widehat{X}$ on Ω but the added cemetery point Δ plays a useful role in the arguments.

Consider the following function $F_{(\xi^m),f}(\cdot)$ defined in (1.93) below from $(\mathcal{M}(E))^{\widehat{\Omega}}$ into $\mathbb{R}$:

$$F_{(\xi^m),f}(\widehat{X}) := \int\limits_E \cdots \int\limits_E f(u_1, u_2, \ldots, u_m)\widehat{x}_{\xi_1}(du_1) \ldots \widehat{x}_{\xi_m}(du_m), \tag{1.93}$$

with

$$\widehat{X} = (\widehat{x}_{\xi})_{\xi \in \widehat{\Omega}}, \quad \xi^m = (\xi_i)_{i=1,\ldots,m} \in \widehat{\Omega}^m, \quad f \in C(E^m), \quad m \in \mathbb{N}. \tag{1.94}$$

For the most part we shall restrict to the subset $(\mathcal{P}(E))^{\widehat{\Omega}}$ of $(\mathcal{M}(E))^{\widehat{\Omega}}$.

Given n different points $\xi_1, \ldots, \xi_n$ in $\widehat{\Omega}$, and multiplicities $n_1, \ldots, n_n$, $m = n_1 + \cdots + n_n$, the *joint moment measure* is obtained by taking expectations of the measure

$$\left(\bigotimes_1^{n_1} \widehat{x}_{\xi_1}\right) \bigotimes \left(\bigotimes_1^{n_2} \widehat{x}_{\xi_2}\right) \cdots \bigotimes \left(\bigotimes_1^{n_n} \widehat{x}_{\xi_n}\right). \tag{1.95}$$

This joint moment measure is determined by the collection $\{E(F_{(\xi^{(m)}),f}(\widehat{X}) : f \in C(E^m)\}$.

Hence if we consider the family

$$\mathcal{F} = \left\{F_{(\xi^m),f} | (\xi_i)_{i=1,\ldots,m} \in \widehat{\Omega}^m, \ \ f = f \in C(E^m), \quad m \in \mathbb{N}\right\}, \tag{1.96}$$

the law of a random element in $(\mathcal{P}(E))^{\widehat{\Omega}}$ is uniquely determined by specifying the expectations of all functions in $\mathcal{F}$. Define for the $(\mathcal{P}(E))^{\widehat{\Omega}}$-valued process $\widehat{X}(t)$,

$$\widetilde{F}_{(\xi^m),f}(t) = E\Big[F_{(\xi^m),f}(\widehat{X}(t))\Big]. \tag{1.97}$$

In order to set up the duality relation we now introduce a bivariate function $H_{f,M}$ and rewrite (1.97) in terms of this function. Namely, for every $f \in C(E)^m$, $m \in \mathbb{N}$, we define the bivariate function, $H_{f,m}$, as

$$\begin{aligned} H_{f,m} \quad &: \quad (\mathcal{P}(E))^{\widehat{\Omega}} \times (\widehat{\Omega})^m \to \mathbb{R} \\ H_{f,m}(\widehat{X}, \eta) \quad &= \quad \int\limits_E \cdots \int\limits_E f(u_1, \ldots, u_m)\, \widehat{x}_{\eta_1}(du_1) \ldots \widehat{x}_{\eta_m}(du_m). \end{aligned} \tag{1.98}$$

Then replacing η above by ξ^m we can now write:

$$\widetilde{F}_{(\xi^m),f}(t) = E\Big[H_{f,m}(\widehat{X}(t), \xi^m)\Big]. \tag{1.99}$$

Using the function $H_{f,n}$ define the duality functions $\mathcal{H}_{f,n}$ on $(\mathcal{P}(E))^{\widehat{\Omega}} \times \mathfrak{Z}_n$ by:

$$\begin{aligned} \mathcal{H}_{f,n}(\widehat{X}, \zeta) \quad &= \quad \mathcal{H}_{f,n}(\widehat{X}, (\eta, \pi)) = \int\limits_E \cdots \int\limits_E f(u_{\pi(1)}, \ldots, u_{\pi(n)})\, \widehat{x}_{\eta_1}(du_1) \ldots \widehat{x}_{\eta_{|\pi|}}(du_{|\pi|}), \\ & \quad\ \eta = (\eta_1, \ldots, \eta_{|\pi|}), \quad |\pi| \leq n, \ \pi \in \Pi_n. \end{aligned} \tag{1.100}$$

(ii) Duality relation

In this subsection we establish the duality relation between the process $(\zeta^T(t))_{t \geq 0}$ of the type defined in subsection 1(b)(ii) and the process $\widehat{X}$, the interacting Fleming-Viot process with time-inhomogeneous and singular rates. This will allow us to determine for fixed time T the value of $\widetilde{F}_{(\xi^m),f}(T)$ (recall (1.97)) by calculating an expectation of the dual process at time T. This is basically a generator calculation plus handling of the singularities as we shall now see.

For $t \le T$, let $\widehat{G}_t$ denote the following operator acting on functions $\widehat{F} \in \mathcal{F}$ (defined in (1.96)):
(α) If $t, \widehat{F}$ are such that $\widehat{F}$ does not depend on any ξ at which $d_\xi(t) = \infty$ or $m_\xi(t) = \infty$, then the operator is:

$$(\widehat{G}_t \widehat{F})(\widehat{X}) = \sum_{\xi \in \Omega} \Bigg[\int_E \frac{\partial \widehat{F}}{\partial \widehat{x}_\xi}(\widehat{X})(u) \Big(\sum_\eta \tilde{a}^t(\xi,\eta)(\widehat{x}_\eta - \widehat{x}_\xi) \Big)(du) + d_\xi(t) \int_E \int_E \frac{\partial^2 \widehat{F}}{\partial \widehat{x}_\xi \partial \widehat{x}_\xi}(\widehat{X})(u,v) Q_{\widehat{x}_\xi}(du,dv) \Bigg], \tag{1.101}$$

where $Q_{\hat{x}_\xi}(du,dv) = \hat{x}_\xi(du)\delta_u(dv) - \hat{x}_\xi(du)\hat{x}_\xi(dv)$.
(β) If t is such that a singularity of the rates $\tilde{a}^t(\xi,\eta)$ or $d_\xi(t)$ occurs at points ξ on which $\widehat{F}$ depends, then we adopt the convention to put $\widehat{G}_t$ equal to the zero operator.

Now fix a time $T > 0$. Given the collection of time-inhomogeneous generators (with singularities in the rates (0.50) - (0.54))

$$(\widehat{G}_t)_{t \in [0,T]} \tag{1.102}$$

we consider a process $(\widehat{X}(t))_{t \ge 0}$ whose law is a solution to the associated martingale problem with singularities given in chapter 0(c).

In the next Proposition we establish a duality relation with respect to the functions $\mathcal{H}_{f,n}$ between this process and the coalescing random walk in random medium $(\zeta^T(t))_{t \ge 0}$ based on the migration rates $\tilde{a}^{T-t}$ and coalescence rates $d_\xi(T-t)$ given in (1.82) and (1.83). This process we constructed in subsection 1(b). Note however that the coalescing random walk starts at time T and runs in reverse time compared to that of the Fleming-Viot process.

Proposition 1.1. *(Duality relation)*
(a) *Let $(\widehat{X}(t))_{t \in [0,T]}$ be any solution to the $(\widehat{G}_t, \widehat{X}_0)$-martingale problem for $t \in [0,T]$ given in (1.101). Let $(\zeta^T(t))_{t \ge 0}$ be the coalescing random walk with a time evolution specified by the rates $(m_\xi(T-t))_{t \in [0,T]}$ and $(d_\xi(T-t))_{t \in [0,T]}$ with ξ running through Ω. The coalescing random walk starts with $\zeta^T(0) = (\eta_1, \ldots, \eta_n;\ \{1\}, \{2\}, \ldots, \{n\})$.*

Assume (i) that $m_\xi(T), d_\xi(T)$ are non-singular at the locations $\xi = \eta_1, \ldots, \eta_n$ and that (ii) the random walk rates are such that none of the processes $\eta_1(t), \ldots, \eta_{\pi(t)}(t)$ hits $\triangle$ for $0 \le t \le T$, almost surely.
Then for all $n \in \mathbb{N},\ f \in C(E^n)$:

$$E\Big[\mathcal{H}_{f,n}(\widehat{X}(T), \zeta^T(0))\Big] = E\ \Big[\mathcal{H}_{f,n}(\widehat{X}_0, \zeta^T(T))\Big]. \tag{1.103}$$

(b) *If we consider the process $\widehat{X}$ extended to a system on $(\mathcal{P}(E))^{\widehat{\Omega}}$, then (1.103) holds even for all initial states ζ_0 with $|\pi_0| < \infty$ and the law of the extended process is uniquely determined.*
(c) *If the $(\hat{G}_t, \hat{X}_0)$-martingale problem arises via the rates induced by (1.82), (1.83) and (1.76), then the assumption in the assertion (ii) in part (a) is satisfied and the duality relation holds for a every $\zeta(0)$ satisfying assumption (i) of part (a).* □

Corollary The law of the Fleming-Viot process in the fluctuating medium is a measurable function of the medium. □

Remark The Proposition tells us that without violating (1.103), we can modify the states of the $\widehat{X}$ process in those components and time points in which the singularities in the rates occur. However if we declare processes differing only in those time-space point as equivalent, then the martingale problem $(\widehat{G}_t, \widehat{X}_0), t \in [0,T]$ has a *unique solution* in the equivalence classes. By considering the extended system we pick a particular member of the equivalence class.

Before beginning the proof we extend the result of Proposition 1.1 to give a representation of the finite dimensional distributions of $\widehat{X}$ using multi-time moments of the form

$$E\left\{\int_0^1 \cdots \int_0^1 f(u_1,\ldots,u_n)\bigotimes_{i=1}^n [\widehat{x}_{\xi_i}(s_i)(du_i)]\right\}. \tag{1.104}$$

Then we obtain an extended duality relation for these expectations by a modification of the coalescing random walk. To do this, we introduce the birth of n_i new particles at the deterministic times $s_1, s_2, \ldots, s_k \in (0,t),\ s_0 = t$ at the locations given by $\eta_0^i,\ i=1,\ldots,k$. Each newborn particle forms a separate partition element. This must be done so that no singularity occurs at an occupied location at times $s_1,\ldots,s_k$. Once the new particles are born the original system of particles and the new particles (and partitions) evolve together following the same mechanism as described above. This process is called $(\zeta^{T,*}(t))_{t\geq 0}$. The population newly born at time s_i is denoted by $\zeta^{T,i}(0)$.

We can easily extend (1.103) to a *space-time duality* as follows:

Proposition 1.2. *Let f^i be an element of $C(E^{n_i})$ and $n_i \in \mathbb{N}, k \in \mathbb{N}$. Then if no singularities in the rates occur at times s_i $i=0,1,\ldots,k$ at the locations of initial, respectively newborn, particles, then:*

$$E\left[\prod_{i=0}^k \mathcal{H}_{f^i,n_i}(\widehat{X}(s_i),\zeta^{T,i}(0))\right] = E\left[\prod_{i=0}^k \mathcal{H}_{f^i,n_i}(\widehat{X}_0,\zeta^{T,*}(T))\right]. \qquad \square \tag{1.105}$$

Proof of Proposition 1.1

Let L_t^T denote the time inhomogeneous generator at time t of the coalescing random walk $(\zeta^T(t))_{0\leq t\leq T}$ defined in Section 1(b)(ii) with rates given by (1.82),(1.83) and started at time 0 in a configuration with locations which are regular at time 0. Recall that we assume (by requiring what we called admissible rates) that the process never leaves the time-space set of regular points. Then the process $(\zeta^T(t))_{t\geq 0}$ can be characterized as the solution of the $(L_t^T, \zeta_0^T)_{t\in[0,T]}$ martingale problem with respect to the test functions $\{\mathcal{H}_{f,n}(\widehat{X}_0,\bullet), \widehat{X}_0 \in (\mathcal{P}(E))^{\widehat{\Omega}},\ C(E^n),\ n\in\mathbb{N}\}$. What we will use in our argument, however, is only that it is a solution, that is, (1.111) below is a local martingale.

For the duality relation we now fix a time T and consider the evolution of the processes in the time interval $[0,T]$ for the original process run forward and for the dual process run backward. Our task is now to verify that the $(L_{T-t}^T)_{t\in[0,T]}$ and $(\widehat{G}_t)_{t\in[0,T]}$ martingale problems satisfy a duality relation as claimed in (1.103).

The key is the following operator equation which is checked by explicit calculation and which holds for all those ζ not covering a singularity at the given time $t\in[0,T]$:

$$\Big(\widehat{G}_t\mathcal{H}_{f,n}(\cdot,\zeta)\Big)(\widehat{X}) = \Big(L_{T-t}^T\mathcal{H}_{f,n}(\widehat{X},\cdot)\Big)(\zeta). \tag{1.106}$$

If we define *independent* versions of $(\widehat{X}(t))_{t\in[0,T]}$ and $(\zeta^T(t))_{t\in[0,T]}$ on one common probability space we can consider for fixed T the quantity (which of course also depends on T, this we suppress):

$$F(t,s) := E(\mathcal{H}_{f,n}(\widehat{X}(t),\zeta^T(s))). \tag{1.107}$$

Then the statement of the proposition reduces to

$$F(T,0) = F(0,T), \tag{1.108}$$

which we are now going to prove. We proceed in seven steps.

Step 1 The key object is a type of semi-martingale decomposition, which we present first.

Fix a function $n \in \mathbb{N}$ and $f \in C(E^n)$. Given $\zeta^T(0)$ we denote by $\mathcal{N}_s$ the set of sites at which a partition element of $\zeta^T(s)$ is located.
Define two stochastic processes $(M_1(s,t))_{0\le s\le t}$ and $(M_2(t))_{t\ge 0}$.

We first define M_1. For $s<t$ and $s,t\in[0,T]$:

$$M_1(s,t) := \mathcal{H}_{f,n}(\widehat{X}(t),\zeta^T(T-s)) - \mathcal{H}_{f,n}(\widehat{X}(s),\zeta^T(T-s)) - \int_s^t h_s(t)[\widehat{G}_r\mathcal{H}_{f,n}](\widehat{X}(r),\zeta^T(T-s))dr, \tag{1.109}$$

where we define

$$h_s(t) = 1_{A(t)}(t), \quad A(t) = \{(d_\xi(r)+m_\xi(r)<\infty, \quad \forall \xi\in\mathcal{N}_s,\ r\in[s,t))\}. \tag{1.110}$$

Next turn to the definition of $(M_2(t))_{t\ge 0}$. We know by assumption that the paths of ζ are such that the locations remain in regular points. Hence we can set for $0\le s<t\le T$,

$$\begin{aligned} M_2(t)-M_2(s) &= \mathcal{H}_{f,n}(\widehat{X}(T-s),\zeta^T(t)) \\ &-\mathcal{H}_{f,n}(\widehat{X}(T-s),\zeta^T(s)) - \int_s^t L^T_{T-r}\mathcal{H}_{f,n}(\widehat{X}(T-s),\zeta^T(r))dr. \end{aligned} \tag{1.111}$$

We next note that according to the martingale problem for $\widehat{X}$ for fixed $\zeta^T(s)$, $(M_1(s,r))_{s\le r\le t}$ is a local martingale with respect to the natural filtration generated by the process $\widehat{X}$ provided that $h_s(t)=1$.

By the martingale problem for ζ^T, $(M_2(r)-M_2(s))_{r\ge s}$ is a local martingale in the time interval $[s,t]$ with respect to the natural filtration generated by the process ζ^T.

We want to now use this structure to obtain a representation of $F(T,0)-F(0,T)$ that will allow us to compute it. The strategy will be to locally freeze one of the processes to exploit the semimartingale structure just described. This leads to a decomposition in terms of differences over small time intervals and we then take the limit as the maximal length of these intervals shrinks to zero. Here is the formal set-up.

Let $\{s_i^k\} := \{0=s_0^k\le s_1^k\le s_2^k\le \dots s_m^k = T\}$ with $m=m(k)$ be a sequence of partitions of $[0,T]$ and let $|s^k| := \max_i |s^k_{i+1}-s^k_i|$. Assume that

$$\lim_{k\to\infty} |s^k| = 0. \tag{1.112}$$

Then using the identity

$$\begin{aligned} &\mathcal{H}_{f,n}(\widehat{X}(0),\zeta^T(T)) - \mathcal{H}_{f,n}(\widehat{X}(T),\zeta^T(0)) \\ &= -\sum_{i=0}^{m-1}\left[\mathcal{H}_{f,n}(\widehat{X}(s^k_{i+1}),\zeta^T(T-s^k_{i+1})) - \mathcal{H}_{f,n}(\widehat{X}(s^n_i),\zeta^T(T-s^k_i))\right], \end{aligned} \tag{1.113}$$

we obtain successively with (1.109) and (1.111) the representation

$$
\begin{aligned}
&F(T,0)-F(0,T)= \\
(1.114)\qquad & E\left[\mathcal{H}_{f,n}(\widehat{X}(T),\zeta^T(0))-\mathcal{H}_{f,n}(\widehat{X}(0),\zeta^T(T))\right] \\
= \;& E\left(\sum_{i=0}^{m-1}\left[\mathcal{H}_{f,n}(\widehat{X}(s^k_{i+1}),\zeta^T(T-s^k_{i+1}))-\mathcal{H}_{f,n}(\widehat{X}(s^k_i),\zeta^T(T-s^k_i))\right]\right) \\
= \;& E\left(\sum_{i=0}^{m-1}\left[\mathcal{H}_{f,n}(\widehat{X}(s^k_{i+1}),\zeta^T(T-s^k_i))-\mathcal{H}_{f,n}(\widehat{X}(s^k_i),\zeta^T(T-s^k_i))\right]\right) \\
&\quad +E\left(\sum_{i=0}^{m-1}\left[\mathcal{H}_{f,n}(\widehat{X}(s^k_{i+1}),\zeta^T(T-s^k_{i+1}))-\mathcal{H}_{f,n}(\widehat{X}(s^k_{i+1}),\zeta^T(T-s^k_i))\right]\right) \\
= \;& E\left(\sum_{i=0}^{m-1}\left(\int_{s^k_i}^{s^k_{i+1}} h_{s^k_i}(s^k_{i+1})\left[\widehat{G}_r\mathcal{H}_{f,n}(\widehat{X}(r),\zeta^T(T-s^k_i))\right]dr\right)\right) \\
+ \;& E\left(\sum_{i=0}^{m-1}\left((1-h_{s^k_i}(s^k_{i+1}))\left[\mathcal{H}_{f,n}(\widehat{X}(s^k_{i+1}),\zeta^T(T-s^k_i))-\mathcal{H}_{f,n}(\widehat{X}(s^k_i),\zeta^T(T-s^k_i))\right]\right)\right) \\
&\quad +E\left(\sum_{i=0}^{m-1} h_{s^k_i}(s^k_{i+1})(M_1(T-s^k_i,T-s^k_{i+1}))\right) \\
&\quad -E\left(\sum_{i=0}^{m-1}\left(\int_{T-s^k_{i+1}}^{T-s^k_i}\left[L_r\mathcal{H}_{f,n}(\widehat{X}(s^k_{i+1}),\zeta^T(r))\right]dr\right)\right) \\
&\quad +E\left(\sum_{i=0}^{m-1}(M_2(T-s^k_i)-M_2(T-s^k_{i+1}))\right).
\end{aligned}
$$

Step 2 In this decomposition one groups the terms into two sets of terms, the integral terms and the differences of the M-processes. For the first group one brings into play the generator relation and for the second martingale properties.

To illustrate the main idea we first consider the case in which there are no singularities. In addition, for the purpose of this discussion we consider the *regular case* which means that the rates $m_\xi(t)$ and $d_\xi(t)$ as functions in t and ξ are bounded on $[0,T]\times\Omega$. In the regular case, by choosing a sequence of partitions with $|s^k|\to 0$ the above expression goes to zero as is seen as follows.

First note that the first and fourth terms above approximate the time integrals and converge to

$$
(1.115)\quad E\left(\int_0^T\left[\widehat{G}_r(\mathcal{H}_{f,n}(\widehat{X}(r),\zeta^T(T-r))\right]dr\right)-E\left(\int_0^T\left[L_{T-r}(\mathcal{H}_{f,n}(\widehat{X}(r),\zeta^T(T-r))\right]dr\right),
$$

which is zero by the operator identity (1.106).

Hence it remains to deal with the expression (here $\tau^k_i=T-s^k_i$ and recall $m=m(k)$)

$$
\begin{aligned}
(1.116)\qquad & \lim_{k\to\infty}\left[\sum_{i=0}^{m-1}E[M_1(\tau^k_{i+1})-M_1(\tau^k_i)]-\sum_{i=0}^{m-1}E[M_2(s^k_{i+1})-M_2(s^k_i)]\right] \\
&=E[(M_1(0,T))+(M_2(T)-M_2(0))].
\end{aligned}
$$

But in the regular case, the local martingales $(M_1(0,t))_{t\geq 0}$ and M_2 are in fact bounded martingales and hence (1.116) is zero and the claim follows.

We now point out in the remaining steps 3-7 how to incorporate the changes needed to handle the *singular case*. What is the problem? First we note that although the process ξ^T does not hit singularities, in the singular case the jump rates for the process ζ^T are unbounded. Moreover, in the singular case the coefficients in the generator $\widehat{G}_t$ are unbounded and discontinuous at the singularities. As a consequence there are two difficulties to overcome. First, we cannot immediately conclude that (1.115) is equal to zero, since the operator identity holds only over time intervals $[s,t]$ not intercepted by a singularity at a location of $\zeta^T(s)$. Second, since the M_1, M_2 are a priori only local martingales in this case, care must be taken in dealing with the martingale increments.

Step 3 In order to circumvent the just described difficulties arising in step 2 for the singular case we first introduce in this step a sequence of *modifications* of the process $(\zeta^T(t))_{t\in[0,T]}$, which will serve as approximate duals and for which the necessary properties can be established. The key point is that these *approximate duals* will give approximations to the moment measures for the process $\widehat{X}$ that converge monotonically.

The modification is based on two ideas: (α) to restrict the dual process to a *finite set* $\Omega^{(m)}$ of active components and the cemetery and (β) to *regularize* in order to keep the time integrals of the rates both for spatial jumps and coalescence events integrable by adding jumps to the cemetery with the part of the total rate which exceeds K once integrated over time. The latter corresponds to a "killing term" for the process $\widehat{X}$ restricted to Ω thus leading to lower bounds for moment measures.

Let $\Omega^{(m)}$ denote an increasing sequence of finite subsets of Ω with $\Omega^{(m)} \uparrow \Omega$ as $m \to \infty$. The sequence of these approximating duals is obtained by replacing all jumps leading from $\Omega^{(m)}$ to its complement by jumps to $\triangle$. Furthermore to avoid large jump rates, we introduce additional jumps to Δ. This way we introduce the type $\tilde{\Delta}$ into the system $\widehat{X}$, since by convention all mass at Δ, $\widehat{x}_\Delta$, has type $\tilde{\Delta}$. Also, recall that all transition rates at Δ are zero (i.e. the mass at Δ is "inert".).

We make this precise as follows. Without loss of generality we can assume that the initial locations of $\zeta^T(0)$ are contained in $\Omega^{(m)}$ and that there are no singularities in $\Omega^{(m)}$ at time T. Let $m \in \mathbb{N}$, $K > 0$. We then define $(\zeta^{(m,K)}(t))_{t\in[0,T]}$ (we now suppress the dependence on T) with $\zeta^{(m,K)}(0) = \zeta^T(0)$. Using the notation of Section 1(b)(ii) we denote this process:

$$(1.117)\qquad (\zeta_t^{(m,K)})_{0\le t\le T} = ((\eta_1(t),\dots,\eta_{|\pi_t|}(t));\pi_t).$$

The process $\zeta^{(m,K)}$ has the same dynamics as ζ^T except for the elimination of jumps to $(\Omega^{(m)})^c \setminus \Delta$ and the addition of jumps to Δ with rates to be specified below using the following quantities. Let

$$(1.118)\qquad R_0^{(m)}(t,\xi) = \sum_{\eta\in(\Omega^{(m)})^c} \frac{m_\xi(t)}{m_\eta(t)} a(\xi,\eta).$$

$$(1.119)\qquad R_1^{(m)}(t,\zeta) = \sum_{i=1}^{|\pi_t|} \sum_{\eta\in\Omega^{(m)}} \frac{m_{\eta_i(t)}(t)}{m_\eta(t)} a(\eta_i(t),\eta)$$

$$(1.120)\qquad R_2^{(m)}(t,\zeta) = \sum_{\xi\in\Omega^{(m)}} \binom{\sum_{i=1}^{|\pi_t|} 1_\xi(\eta_i(t))}{2} d_\xi(t).$$

$$(1.121)\qquad R_t^{(K)}(\zeta^{(m,K)}) := 1\left(\sum_{\ell=1}^2 \int_0^t R_\ell^{(m)}(s,\zeta^{(m,K)}(s))ds > K\right)\left(\sum_{\ell=1}^2 R_\ell^{(m)}(t,\zeta^{(m,K)}(t))\right).$$

Note that $\sum\limits_{\ell=1}^2 (\int_0^t R_\ell^{(m)}(s,\zeta^{(m,K)}(s))ds)$ is a measure of the total activity (internal jumps and coalescence) in $\Omega^{(m)}$ in the time interval $[0,t]$. Note that since ζ^T does not hit singularities or Δ, $\sum_{\ell=1}^2 \left(\int_0^T R_\ell^{(m)}(s,\zeta^T(s))ds\right) < \infty$ a.s. since otherwise there would be infinitely many jumps before T yielding a contradiction. However this rate can have infinite expectation.

Now we are ready to make two modifications of the dynamic $(\zeta^T_t)_{t\in[0,T]}$.

(α) We first eliminate all jumps from ξ to $(\Omega^{(m)})^c$ and add instead jumps of particles at ξ to Δ at the corresponding rate $R^{(m)}_0(t,\xi)$. In other words we replace jumps to $(\Omega^{(m)})^c$ by jumps to Δ.

(β) The second class of jumps to Δ is introduced to ensure that

$$E\left[\sum_{\ell=1}^{2}\left(\int_0^T R^{(m)}_\ell(s,\zeta^{(m,K)}(s))ds\right)\right]<\infty. \tag{1.122}$$

This is done by introducing jumps of all $|\pi_t|$ partition elements at time t to Δ at rate $R^{(K)}_t(\zeta^{(m,K)})$ and suppressing other internal jumps and coalescence if $\int_0^t R^{(m)}_\ell(s,\zeta^{(m,K)}(s))ds>K$. As a result there is at most one additional such jump in the latter case.

This construction given in (α) and (β) above can be modified so that ζ^T and $\zeta^{(m,K)}$ are defined on a *common probability* space in such a way that $\zeta^{(m,K)}$ restricted to $\Omega^{(m)}$ is a subset of ζ^T. We work with this version of the processes in the sequel.

We verify (1.122) as follows. First note that if $r(\cdot)\geq 0$ and $\int_0^t r(s)ds\to\infty$ as $t\to\infty$, then

$$\int_0^\infty r(t)\exp(-\int_0^t r(s)ds)dt<\infty$$

since the integrand is a total derivative of a function which is 1 at 0 and 0 at ∞. This implies that with τ denoting the time of the first jump and if $r(q)$ is the jump rate that:

$$E\left(\int_0^\tau r(q)dq\right)<\infty. \tag{1.123}$$

Hence applying this to our situation the claim (1.122) follows.

This completes the construction of the process $\zeta^{(m,K)}$. Note that $\zeta^{(m,K})$ is not a Markov process, but still can be described in the obvious way by a martingale problem generated by a operators $(L^{m,K}_t)_{t\in[0,T]}$. Note that the martingale problem is given by a modified operator $(L^{m,K}_t)$ that depends on the path $\{\zeta^{(m,K)}(s)\}_{0\leq s\leq t}$.

Step 4 We next apply the modified process $\zeta^{(m.K)}$ to obtain information on the moment measures for $\widehat{X}$. Note that to determine moment measures it suffices to restrict consideration to the collection $\mathcal{H}_{f,n}$ with the additional restrictions on f:

$$(i)\ f\geq 0,\quad (ii)\ f(u_1,\dots,u_n)=0\quad \text{if } u_i=\tilde{\Delta}\quad \text{for some } i. \tag{1.124}$$

Recall the convention $\widehat{X}_\Delta=\delta_{\tilde{\Delta}}$, and that Δ is a trap.

Under these assumptions we verify now for every $r\in[0,T]$ the key *generator inequality*:

$$\widehat{G}_r(\mathcal{H}_{f,n}(\widehat{X}(r),\zeta^{(m,K)}(T-r)))\geq L^{m,K}_{T-r}(\mathcal{H}_{f,n}(\widehat{X}(r),\zeta^{(m,K)}(T-r))), \tag{1.125}$$

provided that $\zeta^{(m,K)}(T-r)$ does not cover a singularity. The inequality (1.125) is the key relation for the argument in the singular case, replacing (1.106) used in the regular case.

To see this first note that if any partition element of $\zeta^{(m,K)}$ is at Δ and $\hat{x}_\Delta=0$, then $\mathcal{H}_{f,n}(\widehat{X}(r),\zeta^{(m,K)}(T-r))=0$. Therefore it suffices to consider the case in which all partition elements are in $\Omega^{(m)}$. Then $L^{m,K}_{T-r}(\mathcal{H}_{f,n}(\widehat{X}(r),\zeta^{(m,K)}(T-r)))$ has the same negative terms as $G_r(\mathcal{H}_{f,n}(\widehat{X}(r),\zeta^{(m,K)}(T-r)))$ but possibly fewer positive terms.

Step 5 Our next task is to draw conclusions from the approximate dual process $\zeta^{m,K}$ for the dual process ζ^T. More precisely define $F^{m,K}$ as F in (1.107) by replacing ζ^T by $\zeta^{(m,K)}$. We now want to prove that

$$F(T,0)-F^{m,K}(0,T) \tag{1.126}$$

is close to 0 for m,K both sufficiently large.

We first introduce some key objects. Let τ_i denote the time of the ith jump of $\zeta^{(m,K)}$ and recall that there are only finitely many jumps in $[0,T]$. For each i let

$$\tau_i^* := \inf\{s > \tau_i, m_\eta(s) = \infty \text{ for some } \eta \in \Omega^{(m)}\},$$

and

$$\kappa := \min\{\tau_i^* - \tau_i\}. \tag{1.127}$$

Note $\kappa > 0$ a.s.. Furthermore on the event that $|s^k| < \kappa$, none of the intervals $[s_i^k, s_{i+1}^k)$ contain a singularity and therefore $h_{s_i^k}(s_{(i+1)}^k) = 1\ \forall i$ on this event.

Step 6 We will next verify that for every m and k

$$F(T,0) - F^{m,K}(0,T) \geq 0. \tag{1.128}$$

Let $\{s^k\}$ be a sequence of partitions of $[0,T]$ with $|s^k| \downarrow 0$. Let N_k denote the number of elements in s^k and define:

$$A_L := \left\{ \int_0^T \sum_{\ell=1,2} R_\ell^{(m)}(s, \zeta^{(m,K)}(s))ds \leq L, \quad \kappa > \frac{1}{L} \right\}. \tag{1.129}$$

Then consider the decomposition arising via a similar calculation as (1.114)

$$\begin{aligned}
&F(T,0) - F^{(m,K)}(0,T) = \\
&E\left(\sum_{i=0}^{N_k-1}\left(\int_{s_i^k}^{s_{i+1}^k} h_{s_i^k}(s_{i+1}^k)\left[\widehat{G}_r \mathcal{H}_{f,n}(\widehat{X}(r), \zeta^{(m,K)}(T-s_i^k))\right] dr\right)\right) \\
+\quad &E\left(\sum_{i=0}^{N_k-1}\left(\int_{s_i^k}^{s_{i+1}^k}(1-h_{s_i^k}(s_{i+1}^k))\right.\right. \\
&\left.\left.\left[\mathcal{H}_{f,n}(\widehat{X}(s_{i+1}^k), \zeta^{(m,K)}(T-s_i^k)) - \mathcal{H}_{f,n}(\widehat{X}(s_i^k), \zeta^{(m,K)}(T-s_i^k))\right] dr\right)\right) \\
&\quad +E\left(\sum_{i=0}^{N_k-1} h_{s_i^k}(s_{i+1}^k)(M_1(T-s_i^k, T-s_{i+1}^k))\right) \\
&\quad -E\left(\sum_{i=0}^{N_k-1}\left(\int_{T-s_{i+1}^k}^{T-s_i^k}\left[L_r \mathcal{H}_{f,n}(\widehat{X}(s_{i+1}^k), \zeta^{(m,K)}(r))\right] dr\right)\right) \\
&\quad +E\left(\sum_{i=0}^{N_k-1}(M_2(s_i^k) - M_2(s_{i+1}^k))\right).
\end{aligned} \tag{1.130}$$

Hence we are left with a sum of five terms each of which we now have to eliminate by choosing the parameters k and L suitably.

First note that

$$\begin{aligned}
\|M_2\|_{var} &:= \sup_k \left(\sum_{i=0}^{m-1} |(M_2(s_i^k) - M_2(s_{i+1}^k))|\right) \\
&\leq \int_0^T \left|L_r \mathcal{H}_{f,n}(\widehat{X}(s_{i+1}^k), \zeta^{(m,K)}(r))\right| dr + const \cdot N^{(m,K)}(T).
\end{aligned}$$

where $N^{(m,K)}(T)$ denotes the number of jumps of $\zeta^{(m,K)}$ in $[0,T]$. By construction $E\|M_2\|_{var} < \infty$ uniformly in k, L. This bound on the expected variation implies (ζ^T is a jump process) that M_2 is a uniformly integrable martingale.

Let $\varepsilon > 0$. Since $P((A_L)^c) \to 0$ as $L \to \infty$, we can choose $L > 0$ such that

$$
\begin{aligned}
E(1_{(A_L)^c}|F(T,0) - F^{(m,K)}(0,T)|) &< \frac{\varepsilon}{3}, \\
E\left(1_{(A_L)^c}|\sum_{i=0}^{N_k-1}(M_2(s_i^k) - M_2(s_{i+1}^k))|\right) &< \frac{\varepsilon}{3}, \\
E\left(\sum_{i=0}^{N_k-1}(M_2(s_i^k) - M_2(s_{i+1}^k))\right) &= 0.
\end{aligned} \tag{1.131}
$$

We can then choose $k_0(L)$ so that $|s^{k_0}| < \frac{1}{L}$. Then on the set A_L the intervals $[\tau_i^k, \tau_i^{k+1})$ do not cover any singularity and therefore by the independence of A_L and M_1,

$$
E\left(1_{(A_L)}\sum_{i=0}^{N_k-1}(M_1(T-s_i^k, T-s_{i+1}^k))\right) = 0 \tag{1.132}
$$

Finally, since all terms are regular on A_L if $k \geq k_0(L)$, we can take the limits $k \to \infty$ and get that the sum of the integral term converges to the quantity on the left side below and by (1.125) satisfies:

$$
\begin{aligned}
& E\left(1_{A_L}\int_0^T \left[(\widehat{G}_r\mathcal{H}_{f,n})(\widehat{X}(r), \zeta^{(m,K)}(T-r))\right] dr\right) \\
& -E\left(1_{A_L}\int_0^T \left[(L_{T-r}^{m,K}\mathcal{H}_{f,n})(\widehat{X}(r), \zeta^{(m,K)}(T-r))\right] dr\right) \geq 0.
\end{aligned} \tag{1.133}
$$

Now we can conclude the argument as follows. By taking $k_0(L)$ sufficiently large we can make the sum of the first and third terms in (1.130) greater than $-\frac{\varepsilon}{3}$. Therefore $F(T,0) - F^{(m,K)}(0,T) \geq -\varepsilon$ which implies (1.128).

Step 7 Next we have to return to the "real" candidate for the dual process ζ^T. Note that $F^{(m,K)}$ is *monotone increasing* in m and K and hence letting $K \to \infty$ and then $m \to \infty$ we get

$$
E[\mathcal{H}_{f,n}(\widehat{X}(0), \zeta^{(m,K)}(T)] \underset{m,K\to\infty}{\longrightarrow} E[\mathcal{H}_{f,n}(\widehat{X}(0), \zeta^T(T))], \quad \text{for all } f \in \mathbb{C}(E^n), n \in \mathbb{N}. \tag{1.134}
$$

Hence we end up with the key inequality:

$$
E\left[\mathcal{H}_{f,n}(\widehat{X}(T), \zeta^T(0))\right] \geq E\left[\mathcal{H}_{f,n}(\widehat{X}(0), \zeta^T(T))\right]. \tag{1.135}
$$

It remains to show that we actually have equality in (1.135).

To carry this out observe that if there is strict inequality for some $0 \leq f_1 \leq 1$ then by applying the same argument to $f_2 := 1 - f_1$ and adding, we obtain a strict inequality for $f \equiv 1$ on $[0,1]$. But this yields a contradiction since there is equality in (1.135) with $f \equiv 1$ on $[0,1]$ since

$$
P\left(\int_0^T \left(\sum_{\ell=1,2} R_\ell^m(s, \zeta^{(m,K)}(s))\right) ds > K\right) \underset{K\to\infty}{\longrightarrow} 0. \tag{1.136}
$$

Using Lemma 1.12, it then follows that for every realization of the total mass process

$$
P(\eta_i^{m,K}(T) \in (\Omega^{(m)})^c \text{ for some } i) \underset{m,K\to\infty}{\longrightarrow} 0. \tag{1.137}
$$

This concludes the proof of the duality relation in the singular case.

2. Construction and uniqueness (Proof of Theorem 1)

In order to prove Theorem 1 we will assume first that $c_\xi(x) \equiv 0$ and then later use a Girsanov argument to get the general case. We shall then prove Theorem 1 in four steps and as a byproduct we get Proposition 0.1, Proposition 0.3 and Proposition 0.4. Here is the strategy and the content of the four steps:

- Verification that the total mass process $(\bar{X}(t))_{t\geq 0}$ can be characterized by a well-posed martingale problem, which we call the $(\bar{G}, \delta_{\bar{X}})$-martingale problem (Proof of Proposition 0.1).
- Verification that the time-inhomogeneous $(\widehat{G}_t, \delta_{\widehat{X}})$-martingale problem with singular rates is well-posed (recall (0.46) and (0.47)). This proves Proposition 0.3 and heavily uses the duality relation from Proposition 1.1.
- Verification that, if we choose the resampling and migration rates for the $(\widehat{G}_t, \delta_{\widehat{X}})$-martingale problem as in (0.57) and (0.58), that is we define the solutions of the $(\bar{G}, \delta_{\bar{X}}), (\widehat{G}_t, \delta_{\widehat{X}})$-martingale problem denoted $\widehat{X}$ and $\bar{X}$ on one probability space, then we get essentially (with respect to a certain equivalence relation) a well-defined bivariate process $(\bar{X}(t), \widehat{X}(t))_{t\geq 0}$.
- This step has two parts. First we show that we construct a solution of the (G, δ_X)-martingale problem, if we define the process X as $X(t) = (\bar{x}_\xi(t)\widehat{x}_\xi(t))_{\xi\in\Omega}$ in terms of on the pair $(\bar{X}, \ \widehat{X})$ obtained in the previous step. Then in the second part we verify that for every solution of the (G, δ_X)-martingale problem, the corresponding relative weights process conditioned on the total mass process solves the $(\widehat{G}_t, \delta_{\widehat{X}})$-martingale problem (Proof of Proposition 0.4).

The main work of this section is in step 4, in particular the second part. For step 3 the the main work was already done in chapter 1, while the other two are fairly standard. Many details of the arguments below will also be used in subsection 4(b) to prove Theorem 3 on the well-posedness of the martingale problem for the historical process. These are therefore spelled out here in considerable detail in this simpler setting.

We give in subsection 2(a) the main concepts, state the needed intermediate results and then prove in subsection 2(b) the key lemma, namely Lemma 2.6 and this requires considerable technical effort. Subsection 2(c) completes the argument.

(a) Well-posedness of the relative weights martingale problem. As announced the argument proceeds in four steps.

Step 1 **($(\bar{G}, \delta_{\bar{X}})$-martingale problem)**

Consider the martingale problem on the space $C([0,\infty), \bar{\mathcal{E}})$ with generator $\bar{G}$ defined on $\mathcal{D}$, where the ingredients are given below (recall (0.5)):

$$\bar{\mathcal{E}} = \{\bar{X} \in ([0,\infty))^\Omega | \ \|\bar{X}\|_\gamma < \infty\}. \tag{2.1}$$

Let $\mathcal{D}$ denote all functions f on $\bar{\mathcal{E}}$ satisfying:

- f depends on finitely many components,
- f is twice continuously differentiable,
- f is bounded, as well as its first and second derivates.

Then define for $f \in \mathcal{D}$ (recall $\sum_\eta a(\xi,\eta)\bar{x}_\eta < \infty$ for $\bar{X} \in \bar{\mathcal{E}}$ by construction of the norm γ):

$$(\bar{G}f)_{(\bar{X})} = \sum_{\xi\in\Omega} \left\{ \sum_{\eta\in\Omega} a(\xi,\eta)(\bar{x}_\eta - \bar{x}_\xi) \frac{\partial f}{\partial \bar{x}_\xi}(\bar{X}) + g(\bar{x}_\xi) \frac{\partial^2 f}{\partial \bar{x}_\xi \partial \bar{x}_\xi}(\bar{X}) \right\}. \tag{2.2}$$

Lemma 2.1. *The $(\bar{G}, \delta_{\bar{X}})$-martingale problem is for $\bar{X} \in \bar{\mathcal{E}}$ well-posed in $C(\mathbb{R}^+, (\mathbb{R})^\Omega)$. The solution is a strong Markov process with paths which take values in $\bar{\mathcal{E}}$.* □

Proof This martingale problem has been studied in the context of interacting diffusions, given by systems of stochastic differential equations for $(\bar{X}(t))_{t\geq 0}$ which was given in (0.37). The strong existence and strong uniqueness of solutions to the corresponding system of SDE's has been established by Shiga and Shimizu [SS] who made moment calculations that prove $L_1(\gamma)(=\bar{\mathcal{E}})$ and $L_2(\gamma)$ are preserved under the evolution. In particular $E(g(\bar{x}_\xi(t)))<\infty$ for starting configurations in $L_2(\gamma)$ and $E(\bar{x}_\xi(t))<\infty$ for initial configurations in $L_1(\gamma)=\bar{\mathcal{E}}$ for every $t\geq 0$. Therefore $(f(\bar{X}(t))-\int_0^t(Gf)(\bar{X}(s))ds)_{t\geq 0}$ is a martingale, respectively local martingale, if we start in $L_2(\gamma)$ respectively $\bar{\mathcal{E}}$.

It is a general fact that this implies the well-posedness of the $(\bar{G},\delta_X)$-martingale problem. (See [RW, 87], vol. 2, V 19 and 20). The continuity of paths and the strong Markov property is also found in [SS].

Now we need to relate the process from the $(\bar{G},\delta_{\bar{X}})$-martingale problem to the functional $\bar{X}(t)$ of our process $(X(t))_{t\geq 0}$, provided there exists a solution to the (G,δ_X)-martingale problem. Namely

Lemma 2.2. *The process $(\bar{X}(t))_{t\geq 0}$, defined in (0.20) as functional of $(X(t))_{t\geq 0}$, solves (if the latter exists) the $(\bar{G},\delta_{\bar{X}})$-martingale problem for every $X\in\mathcal{E}$.* □

Proof Note first that $X(0)\in\mathcal{E}\Longleftrightarrow\bar{X}(0)\in\bar{\mathcal{E}}$. Inserting in the martingale problem for $X(t)$, i.e. (0.13), the function F with $f_i\equiv 1\!\!1$ gives by explicit calculation:

$$(GF)(X)=(\bar{G}\bar{F})(\bar{X}),\quad\text{with }\bar{F}(\bar{X})=\prod_{i=1}^n g_i(\bar{x}_{\xi_i}).\tag{2.3}$$

Hence for $\bar{F}$ the process $(\bar{X}(t))_{t\geq 0}$ satisfies the relation of the martingale problem. Since the algebra generated by functions of the form $\bar{F}$ as above, is a measure determining algebra on $\bar{\mathcal{E}}\subseteq(\mathbb{R}^+)^\Omega$ and $\bar{F}$ is contained in $\mathcal{D}$, we can apply Lemma 2.1 and we are done.

If we combine Lemma 2.1 and Lemma 2.2, we get from the result of [SS] mentioned in the proof of Lemma 2.1 immediately Proposition 0.1.

Step 2 (Well-posedness of the $(\widehat{G}_t,\delta_{\widehat{X}})$-martingale problem)
Here we prove Proposition 0.4 and its corollary. The essential work has been done in chapter 1 by proving Proposition 1.1 (Duality) and by checking in 1(a) that the conditions for Proposition 1.1 are satisfied. The duality relation (1.103) gives the existence of a solution as well as its uniqueness, since the duality functions $\mathcal{H}_{f,n}(\cdot,\zeta)$ determine for every initial state $\widehat{X}\in\mathcal{P}([0,1])^\Omega$ the distribution at a fixed time up to equivalence uniquely, so that for every time t, given the configuration at all times $u\leq s<t$, depends only on the configuration at time s and the transition kernels

$$P_{t,s}=\mathcal{L}(\widehat{X}(t)\in\cdot|\widehat{X}(s)=\widehat{X})\tag{2.4}$$

are $\mathcal{L}(\widehat{X}(s))$-a.s. uniquely prescribed as well as the marginals at all times s. Both these objects determine then uniquely up to equivalence a law for the stochastic process via Kolmogorov's theorem. Any such solution has by what we know from the classical case the path properties required in condition (iii) of Definition (0.4). Since the law is determined up to equivalence only we can choose the one also satisfying the condition (ii) in Definition (0.4) and then obtain a unique object.

Step 3 (Existence and uniqueness of the bivariate $(\widetilde{G},\delta_{(\bar{X},\widehat{X})})$-martingale problem)
We have seen in step 1, that the $(\bar{G}_t,\delta_{\bar{X}})$-martingale problem specifies uniquely a stochastic process and for almost all realizations of that process according to step 2 and Proposition 0.2 (proved in 1(a)) the $(\widehat{G}_t,\delta_{\widehat{X}})$-martingale problem is well-posed. The law of $\widehat{X}$ for fixed path of $\bar{X}$ is a measurable function of that path, as is seen easily from the characterization of the law by the dual process, which shows that moments are measurable functions of the path $\mathcal{L}(\bar{X})$-a.s. (note we can explicitly construct the dual process i.e. the coalescent in a fixed medium from random clocks based on hazard functions containing $\{(\bar{x}_\xi(t))_{t\geq 0},\ \xi\in\Omega\}$). This means that we can define uniquely a transition kernel $K(\bar{X},d\widehat{X})$. Hence we can construct the law of the pair

$$(\bar{X}(t),\widehat{X}(t))_{t\geq 0}.\tag{2.5}$$

By construction of $\widehat{X}$ via duality for given $\bar{X}$ as a Markov process (using only $\bar{X}(s)$ for $s \le t$ for the definition of $\widehat{X}(t)$ and the fact that the paths of the Markov process $\bar{X}$ take values in the $\bar{\mathcal{E}}$-valued functions on $[0,\infty)$ a.s.), this bivariate process is a Markov process in $\bar{\mathcal{E}} \times \mathcal{P}([0,1])^{\Omega}$.

We can regularize the path properties without destroying the Markov property if we introduce in the state space

$$\bar{\mathcal{E}} \times (\mathcal{P}([0,1]))^{\Omega} \tag{2.6}$$

the equivalence relation:

$$(\bar{X},\widehat{X}) \mathrel{\widehat{=}} (\bar{Y},\widehat{Y}) \quad \text{if } \bar{X}=\bar{Y} \quad \text{and } \widehat{x}_\xi = \widehat{y}_\xi \quad \text{for all } \xi \in \Omega \text{ with } \bar{x}_\xi > 0. \tag{2.7}$$

We call the set of equivalence classes $\widetilde{\mathcal{E}}$.

We will now characterize the evolution of this pair in the state space $\tilde{\mathcal{E}}$ by a well-posed martingale problem. Consider all functions $\widetilde{F}$ of the form:

$$\widetilde{F} : (\mathbb{R}^+)^{\Omega} \times (\mathcal{P}([0,1]))^{\Omega} \to \mathbb{R}, \quad \widetilde{F}(\bar{X},\widehat{X}) = \bar{F}(\bar{X})\widehat{F}(\widehat{X}), \tag{2.8}$$

where

$$\widehat{F}(\widehat{X}) = \int f(u_1,\dots,u_n)\widehat{x}_{\xi_1}(du_1)\dots\widehat{x}_{\xi_n}(du_n), \tag{2.9}$$

$$\bar{F}(\bar{X}) = g_1(\bar{x}_{\xi_1})\dots g_n(\bar{x}_{\xi_n}), \tag{2.10}$$

with $n \in \mathbb{N}, \;\; f \in C([0,1]^n), \;\; g_i \in C_b^2([0,\infty))$ and:

$$g_i(0) = 0 \quad \forall i = 1,\dots,n. \tag{2.11}$$

Denote by $\widetilde{\mathcal{A}}$ the algebra of functions generated by functions of the above type. The algebra $\widetilde{\mathcal{A}}$ determines distributions on $\widetilde{\mathcal{E}}$.

Let $\bar{G}$ be as in (2.2). Define $\widehat{G}_{\bar{X}} (= \widehat{G}_{t,\bar{X}})$ to be the Fleming-Viot generator (the $\widehat{G}_t$ of step 2 and Proposition 0.3) with migration rate $\bar{x}_\eta(t)a(\xi,\eta)(\bar{x}_\xi(t))^{-1}$ and the resampling rate $h(\bar{x}_\xi(t))\;(\bar{x}_\xi(t))^{-1}$ at points ξ at time t.

Define now with these two ingredients:

$$(\widetilde{G}\widetilde{F})(\bar{X},\widehat{X}) = \{(\bar{G}\bar{F})(\bar{X})\}\widehat{F}(\widehat{X}) + \{(\widehat{G}_{\bar{X}}(\widehat{F}))(\widehat{X})\}\bar{F}(\bar{X}), \qquad \widetilde{F} \in \widetilde{\mathcal{A}}. \tag{2.12}$$

The expression $\widehat{G}_{\widehat{X}}(\widehat{F})\bar{F}$ is always well-defined, also in points where $\bar{x}_{\xi_i} = 0$ for some $i \in \{1,\dots,n\}$ (recall(2.9)), since $\bar{F}(\bar{X})$ is then 0 and vanishes at the right order (g_i is differentiable at 0). Note in the martingale problem only the term $\int_u^t (\widetilde{G}\widetilde{F})(\bar{X}(s),\widehat{X}(s))ds$ plays a role. The key result of this step 3 is now:

Lemma 2.3. *The $(\widetilde{G},\delta_{(\bar{X},\widehat{X})})$-martingale problem on $C([0,\infty),\widetilde{\mathcal{E}})$ is well-posed for $(\bar{X},\widehat{X}) \in \widetilde{\mathcal{E}}$, i.e. for $\bar{X} \in \bar{\mathcal{E}}$ and $\widehat{X}$ an element of $(\mathcal{P}([0,1]))^{\Omega}$.* $\square$

Proof

Part 1 (Existence)

We have constructed using step 1 and step 2 a candidate for a solution which is given explicitly in (2.5).

To see that this is a solution observe that for every t such that none of the $\bar{x}_{\xi_i}(t), \; i = 1,\dots,n$ are 0, we can explicitly calculate and thereby show that the derivative of $E(\widetilde{F}((\bar{X}(t),\widehat{X}(t)))$ is given by $E[(\widetilde{G}\widetilde{F})(\bar{X}(t),\widehat{X}(t))]$.

To see this denote by $\bar{\mathcal{F}}_t$ the σ-algebra generated by $(\bar{X}(s))_{s\leq t}$ and similarly for $\widetilde{X}$ and $\widehat{X}$. Observe that $(\bar{X}(t+s)-\bar{X}(s))_{s\geq 0}$ is independent of $\widehat{X}(t)$ and then write

$$
\begin{aligned}
E\left[\widetilde{F}(\bar{X}(t+h),\widehat{X}(t+h)) \right. \quad - \quad & \left. F(\bar{X}(t),\widehat{X}(t))\right] = \\
& E[E[(\bar{F}(\bar{X}(t+h))-\bar{F}(\bar{X}(t))|\bar{\mathcal{F}}_t]\widehat{F}(\widehat{X}(t))] \\
& + E[\bar{F}(\bar{X}(t))E[\widehat{F}(\widehat{X}(t+h))-\widehat{F}(\widehat{X}(t))]|\widetilde{\mathcal{F}}_t \vee \bar{\mathcal{F}}_\infty] + o(h)
\end{aligned} \tag{2.13}
$$

and then use the boundedness and smoothness properties of $\widetilde{F}$ and its factors $\bar{F}$ and $\widehat{F}$ to immediately get the assertion. It remains to deal with the t at which 0's occur for the processes $\bar{x}_{\xi_i}(t),\ i=1,\dots,n$. The key point here is, that the combined expression from (2.12) is well-defined and once we integrate the expression over time (recall (1.17)) the value is independent of the convention we choose for $\widehat{X}(t)$ in those points. This gives existence.

Part 2 (Uniqueness)

The uniqueness of the process follows from the fact that $\widetilde{\mathcal{A}}$ is distribution determining on the set $\widetilde{\mathcal{E}}$ and from the following fact we shall see below. Namely we can choose a representative in $\bar{\mathcal{E}}\times(\mathcal{P}([0,1]))^\Omega$ and both the first component and the second component given the first process, must satisfy well-posed martingale problems by construction of $\widetilde{G}$, namely the $(\bar{G},\delta_{\bar{X}})$, $(\widehat{G}_{\bar{X}},\delta_{\bar{X}})$-martingale problem of step 1 and 2. Note that the $\widetilde{G}$ martingale problem is defined with respect to the filtration $\sigma((\bar{X}(s),\widehat{X}(s))_{s\leq t})$ the $\bar{G}$ martingale problem with respect to $\sigma((\bar{X}(s))_{s\leq t})$ and the $\widehat{G}_{\bar{X}}$ one with respect to $\sigma((\widehat{X}(s))_{s\leq t},\ (\bar{X}(t))_{t\geq 0})$.

For the $(\bar{G},\delta_{\bar{x}})$-martingale problem simply use functions $F\in\widetilde{\mathcal{A}}$ of the form (2.8) with $\widehat{F}\equiv 1\!\!1$. Inserting this choice results immediately in the $\bar{G}$ martingale problem.

For the second we would like to pick $\widetilde{F}\in\widetilde{\mathcal{A}}$ of the form (2.8) with $\bar{F}\equiv 1\!\!1$. This does not give a function in $\widetilde{\mathcal{A}}$ unfortunately. However we can use suitable smooth approximations of the following function

$$
\bar{F}(X)=\begin{cases} 1 & \text{if} \quad \bar{x}_{\xi_1},\dots,\bar{x}_{\xi_n}>0 \\ 0 & \text{if} \quad \bar{x}_{\xi_i}=0 \qquad\qquad \text{for some } i\in\{1,\dots,n\}, \end{cases} \tag{2.14}
$$

and which are functions depending only on $\xi_1,\xi_2,\dots,\xi_n$.

Then we condition on $\bar{X}$, to get for our fixed $\widetilde{F}$ for the component $\widehat{F}$, the $\widehat{G}_{\widehat{X}}$-martingale problem *between* singularities relevant for $\widehat{F}$. These singularities are the same as for $\widetilde{F}$ by construction and by the very definition of the martingale problem with singularities. It remains to show the continuity of the laws of the $\widehat{F}(\widehat{X})$ component at times at which singularities occur at one of the sites $\xi_1,\dots,\xi_n$ occurring in the definition of the function $\widehat{F}$ and those singularities satisfy the condition of (0.54). This however is a consequence of the fact that the dual process is well-defined for all times and continuous in law as its starting time approaches the singularity (recall we consider here singularities where an excursion ends or begins). Since the functions $\widehat{F}$ considered here form a probability measure determining algebra we have proved that for the bivariate martingale problem the conditional laws of the second component given the first must satisfy the $(\widehat{G}_{\widehat{X}},\delta_{\widehat{X}})$-martingale problem with respect to the filtration $\sigma((\widehat{X}(s);\bar{X}(s))_{s\leq t})$. Hence for every realization of the first component it satisfies the martingale problem with respect to the filtration $\sigma((\widehat{X}_s)_{s\leq t})$. Note here that $(\bar{X}(s))_{s>t}$ and $(\widehat{X}(s))_{s\leq t}$ are independent conditioned on $\bar{X}(t)$ as is seen from the duality. This completes step 3.

Step 4 (Existence and uniqueness of $(X(t))_{t\geq 0}$)

Now we are at the point, that we can show that the process $(X(t))_{t\geq 0}$ is well-defined by first showing existence in (i) and then uniqueness (the harder part) in (ii).

(i) Construction of a solution

Introduce the following abbreviations:

$$
E=(\mathcal{P}([0,1])\times\mathbb{R}^+)^\Omega \tag{2.15}
$$

$$
E_0=(\mathcal{M}([0,1]))^\Omega \tag{2.16}
$$

and both spaces are equipped with the product topology, where on the components the topology of weak convergence is used. Then E and E_0 are polish spaces.

Next we define a candidate for the solution of our basic (G, δ_X)-martingale problem. First note that for given $X \in \mathcal{E}$ we can uniquely determine an element $\bar{X} \in \bar{\mathcal{E}}$ by setting $\bar{x}_\xi = x_\xi([0,1])$. Furthermore setting $\hat{x}_\xi = x_\xi(\bar{x}_\xi)^{-1}$ for $\bar{x}_\xi > 0$ determines already a unique pair $(\bar{X}, \widehat{X})$ in $\widetilde{\mathcal{E}}$ such that $\bar{x}_\xi \hat{x}_\xi = x_\xi$ (recall the definition of the equivalence relation defined in (2.6) and (2.7)). Define $Z(0) = (\bar{X}(0), \widehat{X}(0))$.

Furthermore we write using the definitions of step 3:

$$Z(t) = (\bar{X}(t), \widehat{X}(t)) \quad t \geq 0. \tag{2.17}$$

Let $\tau : E \to E_0$ be the function:

$$\tau(\bar{X}, \widehat{X}) = (\bar{x}_\xi \widehat{x}_\xi)_{\xi \in \Omega}. \tag{2.18}$$

Note that τ is constant on the equivalence classes and hence maps $\widetilde{\mathcal{E}}$ into $\mathcal{E}$. Then however we can define

$$X(t) = \tau(Z(t)) \quad t \geq 0. \tag{2.19}$$

This is the candidate for solving our martingale problem.

Lemma 2.4. *$(X(t))_{t\geq 0}$ solves the (G, δ_X)-martingale problem for every $X \in \mathcal{E}$.* $\square$

Proof We need to compare the action of $\tilde{G}$ and G on suitable functions. Even though the formal calculations (see Lemma 2.5) are straightforward we need to overcome some technical obstacles.

In order to relate the generators acting on functions on E and E_0, we now consider on E test functions, which are functions of $(\bar{x}_{\xi_1}\widehat{x}_{\xi_1}, \dots, \bar{x}_{\xi_n}\widehat{x}_{\xi_n})$ instead of the pairs $\{(\bar{x}_{\xi_i}, \widehat{x}_{\xi_i}),\ i = 1, \dots, n\}$. In order to carry out this projection, pick a function $F : E_0 \to \mathbb{R}$ of the form

$$F(X) = \int f(u_1, \dots, u_n) x_{\xi_1}(du_1) \dots x_{\xi_n}(du_n) \tag{2.20}$$

with $f \in C^+([0,1]^n),\ n \in \mathbb{N},\ \xi_i \in \Omega$. We then have to argue later on at the end of this point (i) that indeed we can use these test functions rather than the ones used originally in the definition of the process.

These functions F need not have a priori finite expectations under $\mathcal{L}(X_t)$ for X_0 in $\mathcal{E}$. Define for $p > 1$ the space $\mathcal{E}^p$ by replacing (0.5) with $\sum_\xi \bar{x}_\xi^p \gamma(\xi) < \infty$. These spaces are dense in $\mathcal{E}$ (and contained in $\mathcal{E}$).

Furthermore by truncation we can approximate every configuration $x \in \mathcal{E}$ by a configuration in $\mathcal{E}^p$. As long as the expectations of the approximating processes remain finite, we can do generator calculations. See [LS] for such arguments. It is well-known however, that the functions F have for all times t finite expectations, if we start in $\mathcal{E}^p$, with p large enough. This is a property of the system $(\bar{X}(t))_{t\geq 0}$ alone which is well-known and is easily verified with Ito's formula using $g(x) = o(x^2)$ as $x \to \infty$. (Compare [DG1]).

The function F can be written in the form:

$$\begin{aligned} F(X) &= \bar{F}(\bar{X})\widehat{F}(\widehat{X}), \text{ with} \\ \bar{F}(\bar{X}) &= \bar{x}_{\xi_1} \dots \bar{x}_{\xi_n}, \\ \widehat{F}(\widehat{X}) &= \int f(u_1, \dots, u_n) d\hat{x}_{\xi_1}(du_1) \dots \hat{x}_{\xi_n}(du_n). \end{aligned} \tag{2.21}$$

Note that F is not in $\widetilde{\mathcal{A}}$ of step 3, since $\bar{F}(x)$ is not bounded. Nevertheless can we define $\bar{G}\bar{F}$ and $\widehat{G}_{\bar{X}}(\widehat{F})$ for such functions as in (2.21). To continue we need the following identity:

Lemma 2.5. *If $F, \bar{F}$ and $\widehat{F}$ are as in (2.20) and (2.21) then:*

$$(GF)(X) = (\bar{G}\bar{F})(\bar{X})\widehat{F}(\widehat{X}) + (\widehat{G}_{\bar{X}}(\widehat{F}))\bar{F}(\bar{X}). \quad \square \tag{2.22}$$

Proof of Lemma 2.5 The action of G on F can be calculated explicitly as well as the action of $\bar{G}$ on $\bar{F}$ and $\widehat{G}$ on $\widehat{F}$. By inspection the identity (2.22) follows (recall that $F(\bar{X}) = 0$ if $\bar{x}_\xi = 0$ for some i, so that $\widehat{G}_{\bar{X}}(\widehat{F})\bar{F}(\bar{X})$ is well-defined).

Return to the proof of Lemma 2.4. We use, that we can extend the definition of $\widetilde{G}$ to functions $\widetilde{F}$ of the form $(\bar{X}, \widehat{X}) \to \bar{F}(\bar{X})\widehat{F}(\widehat{X})$ with $\bar{F}, \widehat{F}$ of the form (2.20), (2.21) and we find:

$$\widetilde{G}F(\bar{X}, \widehat{X}) = (GF)(X). \tag{2.23}$$

To see this, write $(\bar{G}\bar{F})\widehat{F} + (\widehat{G}_{\widehat{X}}\widehat{F})\bar{F}$ as GF using the concrete form of $\bar{G}, \widehat{G}_{\widehat{X}}$ and G on functions of the form (2.20). Note that GF is well-defined also at time-space points (t, ξ_i) where $\bar{x}_{\xi_i}(t) = 0$.

We see by (2.12) that, since $(\bar{X}(t), \widehat{X}(t))$ solves the $\widetilde{G}$-martingale problem for $\widetilde{F} \in \widetilde{\mathcal{A}}$ and since $\bar{X}$ has finite n-th moments if started in a point in $\bar{\mathcal{E}}^n$, then the process must also satisfy the martingale problem for functions F of the form (2.21) provided $X \in \bigcap_{p>1} \mathcal{E}^p$. This implies that the process $(X(t))_{t\geq 0}$ solves the (G, δ_X)-martingale problem with test functions of the type (2.21) for $X \in \bigcap_{p>1} \mathcal{E}^p$.

This finishes the proof for even more restricted initial states, since the algebra generated by functions of the type of F in (2.21) is measure determining for pairs $(\bar{X}, \widehat{X})$, if we restrict the possible initial states further. For example require that every component $\bar{x}_\xi$ is bounded by a constant. In this case we have uniformly in ξ all exponential moments of the total masses at $t = 0$. Then for all times $t > 0$ we can guarantee that $\bar{X}(t)$ has exponential moments of some order and hence the quantities $E[F(X(t))]$ determine for every $t > 0$ the distribution uniquely.

By a standard truncation and coupling argument for the process $X(t)$, (compare [LS], [G]) namely using approximations $(x_\xi(0) \wedge K)_{\xi\in\Omega}$ to which the above argument is applicable and which for $K \to \infty$ converges over a finite time horizon to the process we want, we finally see that in fact $(X(t))_{t\geq 0}$ is a solution for the (G, δ_X)-martingale problem if $X \in \mathcal{E}$. Namely for the total masses alone this is classical. But we know by the duality that $\widehat{X}(t)$ has the same property and both combine to the desired conclusion. This completes the proof of Lemma 2.4.

(ii) Uniqueness of the martingale problem

Now it remains to show that $\mathcal{L}((X(t))_{t\geq 0})$ is uniquely determined. From step 1 we know that the total mass process is uniquely determined. We defined the relative weight process by putting $\hat{x}_\xi(t) = (\bar{x}_\xi(t))^{-1} x_\xi(t)$, with the convention $0/0 = 0$. Recall the notion of an admissible modification of this process above (0.58) (not to be confused with with the notion of admissible random walk rates). Hence in view of step 2, in particular Lemma 2.1 and 2.2, it suffices to prove:

Lemma 2.6. *Let $(X(t))_{t\geq 0}$ be a solution of the (G, δ_X)-martingale problem with $X \in \mathcal{E}$. Then the process $(\widehat{X}(t))_{t\geq 0}$ has an admissible modification, that satisfies for given total mass process $(\bar{X}(t))_{t\geq 0}$, the $(\widehat{G}_t, \delta_{\widehat{X}})$-martingale problem, in the sense of (0.50) - (0.54) or (1.101).* □

Since the proof is rather long and technical, we devote the next subsection to this purpose.

(b) Proof of the key Lemma 2.6. We first formulate in (2.25) below the key statement, which will be proved later. The proof is then broken into three steps.

First note that $X \in \mathcal{E}$ implies $\bar{X} \in \bar{\mathcal{E}}$ and $\widehat{X}$ is an element of $(\mathcal{P}([0,1]))^\Omega$. Furthermore a process and another admissible modification differ only in the component ξ and times t where $\bar{x}_\xi(t) = 0$. We have to show that the process $\widehat{X}$ must have an admissible modification such that the following holds. If

$$\{P(B|\bar{\mathcal{X}}) \mid B \in \mathcal{B}(\mathcal{P}([0,1])^\Omega);\ \bar{\mathcal{X}} \in C([0,\infty), \bar{\mathcal{E}})\} \tag{2.24}$$

is a regular conditional probability for $\widehat{X}$ given $\bar{X}$, then

$$P(\cdot|\bar{\mathcal{X}}) \text{ satisfies the } (\widehat{G}_t, \delta_{\widehat{X}}) - \text{martingale problem for } \mathcal{L}((\bar{X}(t))_{t\geq 0}) - \text{a.s. all } \bar{\mathcal{X}}. \tag{2.25}$$

First a word concerning the regularity requirements of the $(\widehat{G}_t, \delta_{\widehat{X}})$- martingale problem. Note that the additional requirement at the singularities makes the solution the martingale problem corresponding to the $\widehat{X}(t)$ evolution for given $\{\bar{X}(t),\ t \geq 0\}$ unique, *without* interfering with the original martingale problem for $X(t)$, since at singular time-points the value of $\widehat{x}_\xi(t)$ becomes irrelevant for $x_\xi(t)$ anyway.

Hence we begin below in part (1) of our proof with simply studying the defining expressions appearing in the $(\widehat{G}_t, \delta_{\widehat{X}})$-martingale problem under the given measure $\mathcal{L}((X(t))_{t\geq 0})$ and we reformulate the problem in terms of a more suitable martingale problem based on the increasing process . In part (2) we then verify the assertion for that martingale problem. This part is the core of the argument. Later we will see in part (3) of the proof, how to use this to prove the solution to the $(\widehat{G}_t)$ martingale problem is uniquely characterized by a duality relation together with the regularity requirements in our notions of a solution to the martingale problem.

Step (1): Reformulation

The analysis is based on a reformulation of the $(\widehat{G}_t, \delta_{\widehat{X}})$-martingale problem by another martingale problem based on linear functionals and their increasing processes, which by Ito's lemma is equivalent to the $(\widehat{G}_t, \delta_{\widehat{X}})$-martingale problem, since the latter works with smooth functions which can be approximated locally by quadratic functions. In order to formulate in part (1) the new martingale problem we next introduce a number of objects.

Define for fixed collection $I = I(\widehat{F}) = \{\xi_1, \ldots, \xi_{|I|} \in \Omega\}$ of distinct sites, numbers $\lambda_1, \ldots, \lambda_{|I|} \in \mathbb{R}^+$ and functions $f_i \in C([0,1])$, the function $\widehat{F}$ on $(\mathcal{P}([0,1]))^\Omega$ by:

$$\widehat{F}(\widehat{X}) = \sum_{i\in I} \lambda_i \langle \widehat{x}_{\xi_i}, f_i \rangle. \tag{2.26}$$

We call the points $I(\widehat{F})$ the $\widehat{F}$-*relevant points.*

Introduce corresponding to this the following collection of functions on $(\mathcal{P}([0,1]))^\Omega$:

$$\mathcal{R} = \{\widehat{X} \to \sum_{i=1}^{k} \lambda_i \langle \hat{X}_{\xi_i}, f_i \rangle,\ (\xi_1, \ldots, \xi_k) \in \Omega^k,\ (\lambda_1, \ldots, \lambda_k) \in (\mathbb{Q}^+)^k,\ f_i \in \mathcal{D}^*, \quad \forall \quad i \in \{1, \ldots, k\},\ \ k \in \mathbb{N}\}, \tag{2.27}$$

where $\mathcal{D}^*$ is a fixed *dense and countable* subset of $C([0,1])$. The set $\mathcal{R}$ of functions is countable and suffices as basic set of functions for a martingale problem formulated in terms of local drift and local diffusion function (increasing process).

Then consider for every $\widehat{F} \in \mathcal{R}$ and for every $r \in [0,\infty)$ (we do not display the u and $\widehat{F}$-dependence of M_t):

$$\begin{aligned} M_t - M_r = h_t(r, \widehat{F}) \cdot \Bigg[& \sum_{i\in I} \lambda_i \langle \widehat{x}_{\xi_i}(t), f_i \rangle - \sum_{i\in I} \lambda_i \langle \widehat{x}_{\xi_i}(r), f_i \rangle \\ & - \int_r^t \Big\{ \sum_{i\in I} \lambda_i (\bar{x}_{\xi_i}(s))^{-1} \sum_\eta a(\xi_i, \eta) \bar{x}_\eta(s) (\langle \widehat{x}_\eta(s), f_i \rangle - \langle \widehat{x}_{\xi_i}(s), f_i \rangle) \Big\} ds \Bigg], \qquad t \geq r \end{aligned} \tag{2.28}$$

where again $h_t(s, \widehat{F}) = 1\!\!1(\bar{x}_{\xi_i}(r) > 0 \text{ for } r \in (s,t),\ i \in I)$. Define again for given s the $\bar{X}$-stopping time T by $T = \inf(t > s | h_t(s, \widehat{F}) = 0)$.

For the next step we consider a new martingale problem for the process $(\widehat{X}(t))$ for a fixed realization of $(\bar{X}(t))_{t\geq 0}$ which satisfies Proposition 0.2. This martingale problem, which is equivalent to the $(\widehat{G}_t, \delta_{\widehat{X}})$-martingale problem by a standard argument (ce.g. Lemma 7.2.1 in [D]), is given as follows. We replace the requirement (0.51) by the following.

For every $\widehat{F} \in \mathcal{R}$ the corresponding (by (2.28)) stochastic process satisfies for all pairs r, t in $[0,\infty)$ with $\bar{x}_{\xi_i}(s) \neq 0$ for $r \leq s \leq t \; \forall i \in I(\widehat{F}) = \{1, \dots, n\}$, that

$$(M_s - M_r)_{t\geq s\geq r} \text{ is a square integrable martingale with respect to the filtration } \sigma(\widehat{X}(s), s \leq t) \vee \sigma((\bar{X}(s))_{s\geq 0}) \tag{2.29}$$

and has the increasing process

$$\langle M_s - M_r \rangle_{t\geq s\geq r} = \left(h_s(r,\widehat{F}) \int_r^s \sum_{i=1}^k \lambda_i^2 \left(\frac{h(\bar{x}_{\xi_i}(q))}{\bar{x}_{\xi_i}(q)} \right) \mathrm{Var}_{\,\widehat{x}_{\xi_i(v)}}(f_i) dq \right)_{t\geq s\geq r}. \tag{2.30}$$

The main point of the remaining argument is to overcome the problem of the singularities introduced by the term $(\bar{x}_{\xi_i}(s))^{-1}$. Having established that $(\widehat{X}(t))_{t\geq 0}$ has an admissible modification which satisfies the martingale problem (2.29),(2.30) we then consider the equivalent martingale problem for $\widehat{X}$ conditioned on $(\bar{X}(t))_{t\geq 0}$ obtained by applying the $\widehat{G}$ martingale problem to different functions, namely functions $\widehat{F}$ of the form (1.100). This in turn will be used to get by Proposition 1.1 the following dual characterization (up to admissible modifications) of the conditional law of $\widehat{X}$:

$$E[\mathcal{H}_{f,n}(\widehat{X}(T), \zeta(0))] = E[\mathcal{H}_{f,n}(\widehat{X}(0), \zeta(T))] \tag{2.31}$$

where $\zeta(\cdot)$ is a dual process defined as in Proposition 1.1.

Step (2): Proof of the Martingale Problem Characterization (2.29) and (2.30)
The proof involves a number of technical aspects due to the existence of singularities in both the resampling and migration rates at which the expressions defining the martingale problem are infinite or not well-defined. In order to establish the martingale problem characterization of $\widehat{X}$ over the finite time interval $[0,T]$ it suffices to consider functionals $\widehat{F}$ as defined above based on finite but sufficiently large sets of communicating sites $I(\widehat{F})$. In view of Lemma 1.5 we can assume that there is no simultaneous zero in $I(\widehat{F})$ in the time interval $[0,T]$. The proof is then broken into several parts.

We first give an overview of the main ideas used in handling the singularities. Since we are considering a function $\hat{F}$ that depends on finitely many sites, $\{\xi_i : i = 1, \dots, \ell\}$ the set of singularities is given by $\{t : \bar{x}_{\xi_i}(t) = 0 \text{ for some } i \in \{1, \dots, \ell\}\}$. Then the interval $[0,t]$, can be decomposed into countably many excursions of $\min_{i=1,\dots,\ell} \xi_i(t)$ away from 0. The first key idea of the proof is to find a finite or countable set of random intervals of the form $[\tau_i, \tau_{i+1}]$ such that the random intervals arise from stopping times $\{\tilde{\tau}_i, i \in \mathbb{N}\}$ by pairs of random times, one a stopping time once we look backward, one if we look forward and $\inf_{s\in[\tau_i,\tau_{i+1}]} \bar{x}_{\xi_i}(s) \neq 0$, $i = 1, \dots, \ell$ and to then verify that $\{M_t : t \in [\tau_i, \tau_{i+1}]\}$ is a martingale in these "$\hat{F}$-regular intervals". The next key idea is to show how to construct a sequence of collections of such intervals which exhaust the complement of the set $\{t : \bar{x}_{\xi_i}(t) = 0 \text{ for some } i \in \{1, \dots, \ell\}\}$, that is the set of excursions, in an appropriate sense.

There are two other essential steps. One, carried out in (ii) is the calculation in (2.38) - (2.46) below identifying the migration rates and the resampling rates for $\widehat{X}$ given $\bar{X}$ similar to a result of Perkins ([P2]). The remaining step is to verify the conservation of the martingale property of certain expressions in passing from the filtration $\sigma((\bar{X}(s), \widehat{X}(s))_{s\leq t})$ to $\sigma((\widehat{X}(s))_{s\leq t}, \; (\bar{X}(t))_{t\geq 0})$. This is established in (iv).

(i) Preparation: Time spans with $\widehat{F}$-regular behavior In the next paragraph (ii) we want to use Ito's lemma in order to represent the increments of the process $\widehat{F}(\widehat{X}(t)) = (\sum_{i=1}^{\ell} \lambda_i \langle \widehat{x}_{\xi_i}(t), f_i \rangle)$ between $\widehat{F}$-relevant singularities as a semi-martingale with respect to the filtration

$$\mathcal{F}_s = \sigma((X(r))_{r \leq s}. \tag{2.32}$$

In this set up we are able to apply ordinary stochastic calculus, since we deal with the unconditioned laws and only later we will consider the larger σ-field, containing the complete information on the total mass process. For this we single out the time intervals, where the process $\widehat{F}(X(t))$ behaves nicely. To do this we use the fact that the total mass processes in a finite set can be decomposed into singularity-free open excursion intervals whose complement is of Lebesgue-measure zero. This is the main goal of this point (i).

We now define the random times at which the singularities relevant for $\widehat{F}$ occur. We consider for every $k \in \mathbb{N}$ the sequences (we suppress the k if the term appears as index or argument)

$$T_n = T_n^k, \quad n \in \mathbb{N}, \tag{2.33}$$

of the random time points such that the random intervals (T_{2j}^k, T_{2j+1}^k) describe the excursions away from 0 that reach the value k^{-1} at a time $\tilde{T}_{2j}^k \in (T_{2j}^k, T_{2j+1}^k)$ at some $\xi_i \in I$. These times are defined as follows:

$$\begin{array}{lcl}
T_0^k &=& 0 \\
T_1^k &=& \inf(t| \quad \bar{x}_{\xi_i}(t) = 0 \quad \text{for some } i \in \{1, \dots, \ell\}) \\
\tilde{T}_2^k &=& \inf(t > T_1^k| \quad \bar{x}_{\xi_j}(t) \geq k^{-1} \quad \text{for all } j \in \{1, \dots, \ell\}\}, \\
T_2^k &=& \sup(t | t < \tilde{T}_2^k, \ \bar{x}_{\xi_i}(t) = 0 \quad \text{for some } i \in \{1, \dots, \ell\}) \\
T_3^k &=& \inf(t > \tilde{T}_2^k | \bar{x}_{\xi_i}(t) = 0 \quad \text{for some } i \in \{1, \dots, \ell\}) \\
\tilde{T}_4^k &=& \inf(t > T_3^k | \bar{x}_{\xi_j}(t) \geq k^{-1} \text{ for all } j \in \{1, \dots, \ell\}\}, \\
\vdots &&
\end{array}$$

Note that T_n^k are stopping times for n odd and $\tilde{T}_n^k$ are stopping times for n even but T_n^k are not stopping times for n even. Since $\{t : \bar{x}_{\xi_i}(t) = 0 \quad \text{for some } i \in \{1, \dots, \ell\}\}$ has Lebesgue measure zero the union of these excursions over k exhaust $[0, t]$ for any $t > 0$ except for a set of Lebesgue measure zero.

In order to show that $(M_t)_{t \geq 0}$ of (2.28) has the properties required from the martingale problem, we will consider *for fixed* k first separately for every $n \in \mathbb{N}$ the processes

$$M_t^{(n,k)} := M_t 1\!\!1(T_n^k < t < T_{n+1}^k), t \geq 0, \tag{2.34}$$

which we will then later piece together for the different n, by defining a suitable value in the points $\{T_n^k; k, n \in \mathbb{N}\}$.

The process $M_t^{(n)}$ has singularities in the rates at T_n^k and T_{n+1}^k, which we have to control in order to apply stochastic calculus. We have to distinguish the cases n even and n odd. If n is even the total mass process is strictly positive between T_n^k, T_{n+1}^k and if n is odd and $T_{n+1}^k > T_n^k$ there are many small excursions away from 0 between T_n^k and T_{n+1}^k. It is these latter intervals, which we have to treat by further subdividing them, by making the second parameter k larger. In the limit $k \to \infty$ we then exhaust the whole interval by good ones, i.e. starting at a time which appears as T_n^k with n even for some k. This will be the essential point (vi) of the present part 2.

We start now with $k = 1$ and consider first the case where *n is even.* We next truncate the process in (T_n^k, T_{n+1}^k) on both ends, close to T_{n+1}^k and close to T_n^k and also deal with the fact that T_n^k is not a stopping time when n is even.

From this point until stated otherwise at the end of (iv) we set

$$n \in \mathbb{N} \quad \text{is even.} \tag{2.35}$$

In order to handle the singularity at T_n^k, choose a number $\delta \in (0, 1)$ and then (recall that we have for the moment k fixed) a random variable $\varepsilon = \varepsilon(n, \delta)$, such that $T_n^k + \varepsilon < T_{n+1}^k$. This ε depends on $\{\bar{x}_{\xi_i}(t), \ i = 1, \dots, \ell\}$, but $T_n + \varepsilon$ should be one element of an increasing sequence of $\bar{X}$-stopping times. For this purpose we consider successively the time points where 0 is hit first and then δk^{-1} with $\delta < 1$ is reached first afterwards, that is, the sequence $T_n^{k/\delta}$, $\tilde{T}_n^k$, n odd. This way we obtain now a collection of stopping times and we have to pick out the one falling in our interval of interest. (After having done this choice we don't have a stopping time anymore).

To make things definite, define in case the singularity at T_n occurs at ξ_i for $i \in \mathcal{N} \subseteq \{1, \dots, \ell\}$ the random variable ε by:

$$T_n + \varepsilon = \inf(t > T_n^k | \bar{x}_{\xi_i}(t) \geq \delta k^{-1} \quad \text{for some } i \in \mathcal{N}). \tag{2.36}$$

Define furthermore in order to handle also the singularity at T_{n+1}^k a $\bar{X}$-stopping time:

$$\mathcal{T}_{m,n}^{\varepsilon} = \inf(t > T_n + \varepsilon | \bar{x}_{\xi_i}(t) \leq m^{-1} \text{ for some } i \in \{1, \dots, \ell\}),\ m > k. \tag{2.37}$$

Note that $\mathcal{T}_{m,n}^{\varepsilon} \to T_{n+1}^k$ a.s. as $m \to \infty$ and $\varepsilon(n, \delta, \omega) \to 0$ a.s. as $\delta \to 0$.

(ii) Semimartingale representation in $\widehat{F}$-regular intervals Now comes the crucial calculation which gives the connection between branching and resampling for test functions $\widehat{F}$ of the type (2.26). Recall from Lemma (2.1) that $(\bar{X}(t))_{t \geq 0}$ is a strong Markov process. Define the process $F_m^{n,\varepsilon}(t)$ derived from $\widehat{F}(\widehat{X}(t))$ for fixed n, m and ε as follows:

$$\begin{aligned} F_m^{n,\varepsilon}(t) \quad = \quad & \sum_{i=1}^{k} \lambda_i \langle x_{\xi_i}(t), f_i \rangle \langle x_{\xi_i}(t), 1 \rangle^{-1}\ 1\!\!\mathrm{I}(T_n + \varepsilon \leq t < \mathcal{T}_{m,n}^{\varepsilon}) \\ & + \sum_{i=1}^{k} \lambda_i \langle x_{\xi_i}(\mathcal{T}_{m,n}^{\varepsilon}), f_i \rangle \langle x_{\xi_i}(\mathcal{T}_{m,n}^{\varepsilon}), 1 \rangle^{-1} 1\!\!\mathrm{I}(t \geq \mathcal{T}_{m,n}^{\varepsilon} > T_n + \varepsilon) \\ & + \sum_{i=1}^{k} \lambda_i \langle x_{\xi_i}(T_n + \varepsilon), f_i \rangle \langle x_{\xi_i}(T_n + \varepsilon), 1 \rangle^{-1} 1\!\!\mathrm{I}(t \leq T_n + \varepsilon). \end{aligned} \tag{2.38}$$

For this process $F_n^{n,\varepsilon}$ we shall derive a *semimartingale representation* for $t \in [0, \infty)$.

As a first step we introduce below in (2.40) explicitly a *stochastic integral* $(M_t^{(n), \mathcal{T}_{m,n}^{\varepsilon}})_{t \geq 0}$. Denote the martingale parts of the semi-martingales $\langle x_{\xi_i}(t), f_i \rangle$ and $\langle x_{\xi_i}(t), 1 \rangle$ by:

$$N^i(t) \quad \text{and } L^i(t). \tag{2.39}$$

Define (recall the discussion below (2.35)):

$$M_t^{(n), \mathcal{T}_{m,n}^{\varepsilon}} = \int_{T_n + \varepsilon}^{t \wedge \mathcal{T}_{m,n}^{\varepsilon}} \sum_{i=1}^{\ell} \lambda_i (\bar{x}_{\xi_i}(s))^{-1} dN^i(s) - \int_{T_n + \varepsilon}^{t \wedge \mathcal{T}_{m,n}^{\varepsilon}} \sum_{i=1}^{\ell} \lambda_i \langle x_{\xi_i}(s), f_i \rangle (\bar{x}_{\xi}(s))^{-2} dL^i(s). \tag{2.40}$$

The key point of this step is, that Ito's formula applied to $\langle x_{\xi_i}, f_i \rangle \langle x_{\xi_i}, 1 \rangle^{-1}$ gives the following two properties:

$$\begin{aligned} F_m^{n,\varepsilon}(t) = {} & F_m^{\varepsilon}(T_n + \varepsilon) + M_t^{(n), \mathcal{T}_{m,n}^{\varepsilon}} \\ & + \int_{T_n + \varepsilon}^{t \wedge \mathcal{T}_{m,n}^{\varepsilon}} \left[\sum_{i=1}^{\ell} \lambda_i \sum_{\eta} \left\{ a(\xi_i, \eta) \left(\frac{\langle x_\eta(s), f_i \rangle - \langle x_{\xi_i}(s), f_i \rangle}{\langle x_{\xi_i}(s), 1 \rangle} \right) \right. \right. \\ & \left. \left. - a(\xi_i, \eta) \frac{\langle x_\eta(s), 1 \rangle - \langle x_{\xi_i}(s), 1 \rangle}{\langle x_{\xi_i}(s), 1 \rangle^2} \langle x_{\xi_i}(s), f_i \rangle \right\} \right] ds \\ = {} & F_m^{n,\varepsilon}(T_n + \varepsilon) + M_t^{(n), \mathcal{T}_{m,n}^{\varepsilon}} + \int_{T_n + \varepsilon}^{t \wedge \mathcal{T}_{m,n}^{\varepsilon}} \sum_{i=1}^{\ell} \sum_{\eta} \lambda_i a(\xi_i, \eta) \frac{\bar{x}_\eta(s)}{\bar{x}_\xi(s)} (\langle \widehat{x}_\eta(s), f_i \rangle - \langle \widehat{x}_{\xi_i}(s), f_i \rangle) ds. \end{aligned} \tag{2.41}$$

The quadratic variation process of the stochastic integral $M_t^{(n),\mathcal{T}^\varepsilon_{m,n}}$ is:

$$\langle M_t^{(n),\mathcal{T}^\varepsilon_{m,n}}\rangle = \left(2\int_{T_n+\varepsilon}^{t\wedge\mathcal{T}^\varepsilon_{m,n}} \sum_{i=1}^{\ell}\left\{\lambda_i^2(\langle\widehat{x}_{\xi_i}(s),f_i^2\rangle - \langle\widehat{x}_{\xi_i}(s),f_i\rangle^2)\right\}\left\{h(\bar{x}_{\xi_i}(s))(\bar{x}_{\xi_i}(s))^{-1}\right\}ds\right)_{t\geq 0}. \tag{2.42}$$

The clue in the proof of (2.41) is the fact that for a function F on $\mathcal{M}([0,1])$ given in the form $F(x) = \langle x,f\rangle\langle x,1\rangle^{-1}$ one has:

$$\int\int \frac{\partial^2 F}{\partial x^2}[u,v]h(\bar{x})x(du)\delta(dv) = 0. \tag{2.43}$$

Namely the second derivative of F reduces due to the linearity of $\langle x,f\rangle$ and $\langle x,1\rangle$ to

$$-\frac{f(u)1(v)}{\langle x,1\rangle^2} - \frac{f(v)1(u)}{\langle x,1\rangle^2} + 2\frac{1(v)1(u)\langle x,f\rangle}{\langle x,1\rangle^3}, \tag{2.44}$$

which in turn after integration with the diffusion operator

$$x(du)\delta_u(dv), \tag{2.45}$$

gives 0. Hence it suffices to calculate $\frac{\partial F}{\partial x}$ and use (2.39) in order to get with the Ito-formula the drift terms as given in (2.41).

Next turn to (2.42). The quadratic variation of the stochastic integrals defined in (2.40) is in view of (2.41) found by calculating the quadratic variation of $F_m^{n,\varepsilon}(t)$. Hence consider again the function $F(x) = \langle x,f\rangle\,\langle x,1\rangle^{-1}$ and the action of the second order part of the generator on the square of this function, which gives

$$\int\int \frac{\partial^2}{\partial x}\left(\frac{\langle x,f\rangle}{\langle x,1\rangle}\right)^2[u,v]h(\bar{x})x(du)\delta_u(dv). \tag{2.46}$$

Then use the chain rule on the square and then (2.43) to get finally formula (2.42).

(iii) The limit $m\to\infty$. First we focus on the right boundary of the interval of regularity, and we will consider $F_m^{n,\varepsilon}(t)$ as $m\to\infty$ to obtain the process $F^{n,\varepsilon}(t)$ and the representation of the new process $F^{n,\varepsilon}(t)$ as a stochastic integral in the interval $[T_n+\varepsilon, T_{n+1})$.

We study now $M_t^{(n),\mathcal{T}^\varepsilon_{m,n}}$ as a function of m. Fix now $n\in\mathbb{N}$. Denote by D_ℓ^T, with $\ell\in\mathbb{N}\cup\{0\}$ the event (for some fixed $T\in\mathbb{R}^+$ and the convention $\mathcal{T}_0^\varepsilon = T_n+\varepsilon$)

$$D_\ell^T = \left\{T_n+\varepsilon+T > \mathcal{T}^\varepsilon_{\ell-1}\right\}\cap\left\{T_n+\varepsilon+T\leq\mathcal{T}^\varepsilon_\ell\right\},\quad D_\infty^T = \left\{T_n+\varepsilon+T>\mathcal{T}^\varepsilon_\ell,\quad\forall\ell\in\mathbb{N}\right\}. \tag{2.47}$$

Note that

$$D_\infty^T = \{T_n+\varepsilon+T\geq T_{n+1}\}. \tag{2.48}$$

We have the following situation on the events D_ℓ^T for $\ell\in\mathbb{N}$ and $m>\ell$. Since $M_s^{(n),\mathcal{T}^\varepsilon_{m,n}}$ agrees for $s\in[T_n+\varepsilon,\ T_n+\varepsilon+T]$ with $M_s^{(n),\mathcal{T}^\varepsilon_\ell}$, we see by the definition of D_ℓ^T, that we can pick a subsequence such that we get Cauchy-series even in the function space $C[T_n+\varepsilon, T_n+\varepsilon+T],\mathbb{R})$ on the event D_ℓ^T. Since we can do this on every of the countably many events D_ℓ^T, $\ell\in\mathbb{N}$ we can define properly a limit for fixed T on the complement of the event $T\geq T_{n+1}-(T_n+\varepsilon)$.

Furthermore if we consider T in a countable set (for example $\mathbb{Q}$), we can piece things together in a consistent way. This defines then a path $(\widetilde{M}^{(n),\varepsilon}_{T_n+\varepsilon+s})_{s\in[0,T_{n+1}-(T_n+\varepsilon))}$ in the limit $m\to\infty$. We define $M_t^{(n),\varepsilon}$ separately on the events that t falls into the interval, to the left or to the right.

$$
\begin{aligned}
t \in [T_n+\varepsilon,\ T_{n+1}) \quad &: M_t^{(n),\varepsilon} = \widetilde{M}^{(n),\varepsilon}_{T_n+\varepsilon+s}, \qquad s = t-(T_n+\varepsilon) \\
t \in [T_{n+1},\infty) \quad &: M_t^{(n),\varepsilon} = 0 \\
t \in [T_n, T_n+\varepsilon] \quad &: M_t^{(n),\varepsilon} = M^{(n)}_{T_n+\varepsilon} = 0.
\end{aligned} \tag{2.49}
$$

This gives a continuous process $(M_t^{(n),\varepsilon})$ in t for which we have for every m:

$$
M_t^{(n),\varepsilon} = M_t^{(n),\mathcal{T}^\varepsilon_{m,n}} \quad \text{on the event } t \in [T_n+\varepsilon,\ \mathcal{T}^\varepsilon_{m,n}]. \tag{2.50}
$$

(Note that at this stage the T_n, ε etc. are still random variables since we have not yet conditioned on the complete total mass process.)

By (2.49) we have the representation:

$$
\begin{aligned}
\sum_{i=1}^{\ell} \lambda_i \langle \widehat{x}_{\xi_i}(t), f\rangle &= \sum_{i=1}^{\ell} \lambda_i \langle \widehat{x}_{\xi_i}(T_n+\varepsilon), f_i\rangle \\
&+ \int\limits_{T_n+\varepsilon}^{t} \left\{ \sum_{i=1}^{\ell} \lambda_i (\bar{x}_{\xi_i}(s))^{-1} \sum_{\eta} a(\xi,\eta)\bar{x}_\eta(s)(\langle \widehat{x}_\eta(s), f_i\rangle - \langle \widehat{x}_{\xi_i}(s), f_i\rangle) \right\} ds \\
&+ M_t^{(n),\varepsilon}, \quad \text{on the event } t \in [T_n+\varepsilon,\ T_{n+1}).
\end{aligned} \tag{2.51}
$$

With this representation we can analyze $(M_t^{(n),\varepsilon})_{t\geq 0}$ further in the next step (iv) and then in (v) we can let $\varepsilon = \varepsilon_n \to 0$.

(iv) The martingale property of $(M_t^{(n),\varepsilon})_{t\geq 0}$ for $\mathcal{F}_t \vee \sigma(\bar{X}(s))_{s\geq 0})$ It now remains to check that $(M_t^{(n),\varepsilon})$ is actually (between the singularities) a martingale with respect to the larger σ-algebra containing the information on all of $\bar{X}(t)$ that is:

$$
\mathcal{H}_t = \sigma((\widehat{X}(s))_{s\leq t}) \vee \sigma((\bar{X}(s))_{s\geq 0}). \tag{2.52}
$$

Note first that now the random variable $\varepsilon = \varepsilon(n,\delta)$ becomes deterministic.

We shall verify in this paragraph that:

Lemma 2.7. *For every bounded and $\sigma((\bar{X}(t))_{t\geq 0})$-measurable random variable R the increments of $M^{(n),\varepsilon}$ are orthogonal to R that is for all $t \geq s$:*

$$
E[(M_t^{(n),\varepsilon} - M_s^{(n),\varepsilon})R|\mathcal{F}_s] = 0, \tag{2.53}
$$

where $\mathcal{F}_s$ is the filtration associated with $(X(r))_{r\leq s}$. □

This will imply the martingale property of $(M_t^{(n),\varepsilon})_{t\geq 0}$ between singularities.

Proof of Lemma 2.7 For (2.53) it suffices to verify this relation for random variables, which depend on only finitely many of the component-processes $\{(\bar{x}_\xi(t))_{t\geq 0},\ \xi \in \Omega\}$. We denote the collection of such components of $(\bar{X}(t))_{t\geq 0}$ associated with R by $\mathcal{I}_R$ and once we have fixed R in a calculation we shall just write $\mathcal{I}$. The verification of (2.53) will proceed by approximation via the $(M_t^{(n),\mathcal{T}^\varepsilon_{m,n}})$ of (2.40).

We shall use a representation of R in terms of a sum of stochastic integrals of the total mass processes. A useful tool for that is the following fact about the total mass processes, which follows immediately from (2.2):

$$
\left(\bar{x}_\xi(t) - \int\limits_0^t \sum_\eta a(\xi,\eta)(\bar{x}_\eta(s) - \bar{x}_\xi(s))ds \right)_{t\geq 0}, \quad \xi \in \Omega \tag{2.54}
$$

are continuous martingales with respect to the σ-algebra $\sigma((\bar{X}(s))_{s\leq t})$. We abbreviate these martingales now by

$$\{(m_\xi(t))_{t\geq 0},\ \xi\in\Omega\}. \tag{2.55}$$

If they are square integrable (for example if $X(0)\in\mathcal{E}^2$), then their increasing processes are given by:

$$\left(\int_0^t h(\bar{x}_\xi(s))\bar{x}_\xi(s)ds\right)_{t\geq 0}. \tag{2.56}$$

Due to (2.54) we have the following representation for the random variable R (see [RY], Proposition V.3.2) for the one-dimensional version). We prove this at the end of (iv).

Lemma 2.8. *There exist processes $(r_\eta(t))_{t\geq 0}$, which are $\sigma(\{\bar{x}_\eta(s), s\geq 0,\ \eta\in I\})$-predictable such that:*

$$R=ER+\sum_{\eta\in I}\int_0^\infty r_\eta(s)dm_\eta(s). \qquad \Box \tag{2.57}$$

Note first that in order to check (2.53) for the approximating increments of $M^{(n),\mathcal{T}^\varepsilon_{m,n}}$, it suffices to consider in (2.57) the integral from s to t. Namely the part from 0 to s is $\mathcal{F}_s$ measurable and we get a zero contribution by the martingale difference property of $(M_t^{(n),\mathcal{T}^\varepsilon_{m,n}})$ expressed in (2.40). Similarly the contribution from t to ∞ is shown to be 0 by first conditioning on $\mathcal{F}_t$ and then using that $m_\eta(r)-m_\eta(t)$ for $r>t$ is independent of $\mathcal{F}_t$ and has mean zero. For the same reasoning no contribution arises from $t\wedge\mathcal{T}^\varepsilon_{m,n}$ till t.

Next observe that now we are able to calculate with (2.40) and (2.57) as follows:

$$\begin{aligned}
&E\left[\left(M_t^{(n),\mathcal{T}^\varepsilon_{m,n}}-M_s^{(n),\mathcal{T}^\varepsilon_{m,n}}\right)R\,\middle|\,\mathcal{F}_s\right] \\
&= E\Bigg[\Bigg\{\int_s^{t\wedge\mathcal{T}^\varepsilon_{m,n}}\sum_{i=1}^{\ell}\lambda_i(\bar{x}_{\xi_i}(r))^{-1}dN^i(r)-\int_s^{t\wedge\mathcal{T}^\varepsilon_{m,n}}\sum_{i=1}^{\ell}\lambda_i\langle x_{\xi_i}(r),f_i\rangle(\bar{x}_{\xi_i}(s))^{-2}dL^i(r)\Bigg\} \\
&\qquad\qquad \Bigg\{\sum_{\eta\in I}\int_s^{t\wedge\mathcal{T}^\varepsilon_{m,n}} r_\eta(r)dm_\eta(r)\Bigg\}\Bigg|\mathcal{F}_s\Bigg].
\end{aligned} \tag{2.58}$$

In order to evaluate the above expressions we have to consider the variations, respectively covariations of the martingales $N_t^i,\ L_t^i$ with $m_\eta(t)$ (recall (2.39) and note that $dm_{\xi_i}(t)=dL^i(t)$ by construction). In order to be able to perform L_2-calculations first assume $X(0)\in\mathcal{E}^2$ so that $E[\bar{x}_\xi(t)]^2<\infty$ for all $t>0$. Then use that by definition of $X(t)$ one must have (calculate the action of the generator G of $(X(t))$ on the functions $\langle x_\xi,f\rangle\langle x_\xi,1\rangle,\ \langle x_\xi,f\rangle\langle x_\eta,g\rangle,\ \langle x_\xi,1\rangle\ \langle x_\eta,1\rangle$ and read of the Ito terms):

$$\begin{array}{llllll}
dL^idL^i &=& \langle x_{\xi_i},1\rangle h(\bar{x}_{\xi_i})dr & dm_{\xi_i}dm_\eta &=& 0\quad \xi_i\neq\eta \\
dL^idL^j &=& 0\quad i\neq j & dN^idN^j &=& 0\quad i\neq j \\
dN^idN^i &=& \langle x_{\xi_i},f_i^2\rangle\ h(\bar{x}_{\xi_i})dr & dN^idL^i &=& \langle x_{\xi_i},f_i\rangle h(\bar{x}_{\xi_i}) \\
dN^idN^j &=& 0\quad i\neq j. & & &
\end{array} \tag{2.59}$$

If we use (2.59) in (2.58) we see that the r.h.s. is equal to

$$E\left[\int_s^{t\wedge\mathcal{T}^\varepsilon_{m,n}}\sum_{i=1}^{\ell}\lambda_i\left\{(\bar{x}_{\xi_i}(r))^{-1}\langle x_{\xi_i}(r),f_i\rangle h(\bar{x}_{\xi_i}(r))-(\bar{x}_{\xi_i}(r))^{-1}\langle x_{\xi_i}(r),f_i\rangle h(\bar{x}_{\xi_i}(r))\right\}dr\Big|\mathcal{F}_s\right]=0. \tag{2.60}$$

Then let $m \to \infty$ to get the desired conclusion (2.53) for $X(0) \in \mathcal{E}^2$. The result is easily generalized to $X(0) \in \mathcal{E}$, by approximation with truncated states $\{x_\xi(0) \wedge n\}_{\xi \in \Omega}$ with $n \to \infty$. We leave the details to the reader. In order to complete the proof we have to verify Lemma 2.8.

Proof of Lemma 2.8 Define $\widehat{R} = R - E(R)$. Then we can define

$$\widehat{R}_t = E(\widehat{R} | \sigma(\bar{X}(s))_{s \le t}). \tag{2.61}$$

Then the process $(\widehat{R}_t)_{t \ge 0}$ has a version which is for $X(0) \in \mathcal{E}^2$ a square integrable closed martingale.

Next observe that the $\{dm_\eta(t),\ \eta \in \Omega\}$ are a orthogonal collection of martingales since they are driven by independent Brownian motions. Now we can apply the multidimensional martingale representation theorem ([KS], Theorem 4.15) to get progressively measurable processes $s_\eta(s))_{s \ge 0}$, $\eta \in I$ such that

$$\widehat{R}_t = \sum_{\eta \in I} \int_0^t s_\eta(q) dw_\eta(q). \tag{2.62}$$

We can now use

$$m_\eta(t) = \int_0^t \sqrt{g(\bar{x}_\eta(q))} dw_\eta(q) \tag{2.63}$$

to obtain

$$\widehat{R}_t = \sum_{\eta \in I} \int_0^t \frac{s_\eta(q)}{\sqrt{g(\bar{x}_\eta(q))}} dm_\eta(q), \tag{2.64}$$

with the convention $\frac{0}{0} = 0$:

$$r_\eta(q) = s_\eta(q) / \sqrt{g(\bar{x}_\eta(q))}. \tag{2.65}$$

This completes the proof of Lemma 2.8.

(v) The limit $\delta \to 0$. We finally come to the left boundary of the interval of regularity for $\widehat{F}$. Here we need a small modification of $(M_t^{(n),\varepsilon})$ in order to be able to vary ε. Recall that $\varepsilon = \varepsilon(n, \delta, \omega)$. We change $M_t^{(n),\varepsilon}$ by a suitable *constant in* t as follows:

$$\widetilde{M}_t^{(n),\varepsilon} = M_t^{(n),\varepsilon} + \Delta(\varepsilon, \varepsilon_0) \tag{2.66}$$

with $\Delta(\varepsilon, \varepsilon_0)$ defined for $0 < \varepsilon < \varepsilon_0$ as follows. Let $\varepsilon_0 = \tilde{T}_n^k - T_n^k$ (recall (2.34)). Then:

$$\Delta(\varepsilon, \varepsilon') = \left\{ \widehat{F}(T_n^k + \varepsilon) - \widehat{F}(T_n^k + \varepsilon') + \int_{T_n^k + \varepsilon}^{T_n^k + \varepsilon'} \sum_{i=1}^{\ell} \sum_\eta a^t(\xi_i, \eta) \left(\widehat{x}_\eta(s) - \widehat{x}_{\xi_i}(s)\right) ds \right\}. \tag{2.67}$$

Note that then the representation of $\sum_{i=1}^n \lambda_i \langle \widehat{x}_\xi(t), f \rangle$ in (2.51) implies for the $(\widetilde{M}_t^{(n),\varepsilon})$, that for $\delta_1 < \delta_2$ one has that $\varepsilon_1 < \varepsilon_2$ a.s. Hence

$$\widetilde{M}_t^{(n),\varepsilon_1} = \widetilde{M}_t^{(n),\varepsilon_2} \quad t \in [T_n^k + \varepsilon_2, T_{n+1}^k), \text{a.s.} \tag{2.68}$$

Therefore we can immediately conclude by letting δ run through the sequence j^{-1} that $\widetilde{M}_t^{(n),\varepsilon}$ converges as $\delta \to 0$ to a limit $M_t^{(n)}$, uniformly for every closed set contained in (T_n^k, T_{n+1}^k). Furthermore:

$$M_t^{(n)} = \widetilde{M}_t^{(n),\varepsilon} \quad \text{for } t \in [T_n^k + \varepsilon, T_{n+1}^k). \tag{2.69}$$

So far we assumed that n is even. We now define

$$M_t^{(n)} = 0 \quad \text{for } t \in (T_n^k, T_{n+1}^k], \text{ with } n \text{ odd.} \tag{2.70}$$

Having (2.53), (2.67) at hand $\{(M_t^{(n)}),\ t \in (T_n^k, T_{n+1}^k)\}$ is then for every $n \in \mathbb{N}$ a process, which is the sum of a random variable plus a local $\mathcal{H}_t$-martingale with continuous path and for every $t \in (T_n^k,\ T_{n+1}^k)$ *with* n *even*, the following holds for every $\delta > 0$:

$$\begin{aligned}\sum_{i=1}^{\ell} \lambda_i \langle \widehat{x}_{\xi_i}(t), f_i\rangle &= \sum_{i=1}^{\ell} \lambda_i \langle \widehat{x}_{\xi_i}(T_n^k + \delta), f_i\rangle \\ &+ \int_{T_n^k+\delta}^{t} \sum_{i=1}^{\ell} \sum_{\eta} a(\xi_i, \eta) \langle (\widehat{x}_\eta(s) - \widehat{x}_{\xi_i}(s)), f_i\rangle \ \frac{\bar{x}_\eta(s)}{\bar{x}_{\xi_i}(s)} ds + M_t^{(n)} + \Delta(\delta, \varepsilon_0), \\ & \qquad \forall \ \ 0 < \delta + T_n^k < t,\end{aligned} \tag{2.71}$$

$$\langle M_t^{(n)} \rangle_{t \geq T_n + \delta} = 2 \int_{T_n^k+\delta}^{t} \sum_{i=1}^{\ell} \left\{ \lambda_i^2 \operatorname{Var}_{\widehat{x}_{\xi_i}(r)} (f_i) \cdot h(\bar{x}_{\xi_i}(r)) (\bar{x}_{\xi_i}(r))^{-1} \right\} dr. \tag{2.72}$$

The process $M^{(n)}$ constructed by piecing together the martingales over the excursions that reach $\frac{1}{k}$ will still *depend on the parameter* k and this process now characterizes M_t in the intervals (T_n^k, T_{n+1}^k) for n even. However, for n odd the definition of $M_t^{(n)}$ is just a temporary convention to get a process everywhere well-defined. In order to complete the process , we will next incorporate all excursions by letting $k \to \infty$ (recall (2.34)).

(vi) *The limit* $k \to \infty$. What we have done to this point is

- to identify the martingale problem (with respect to the larger filtration $\{\mathcal{H}_t\}_{t \geq 0}$) satisfied by the system $\{\widehat{x}_{\xi_i}(t) : i \in I\}$ for t belonging to a union of regular stochastic intervals of the form, $(T_{2n}^k + \varepsilon, \mathcal{T}_{m,n}^{\varepsilon})$, that is, intervals with $\{\mathcal{F}_t\}$ stopping time end points and on which the $\bar{x}_\xi(u) \geq \frac{\delta}{k} \wedge \frac{1}{m}$ and contained within "excursions away from 0 that reach the height $\frac{1}{k}$".
- for each k as the parameters $m \to \infty,\ \delta \to 0$ these stochastic intervals expand to exhaust each excursion that reaches height $\frac{1}{k}$, denoted above by $\Delta_n = (T_{2n}^k, T_{2n+1}^k)$.

To complete the identification of the martingale problem (2.29),(2.30) satisfied by $\widehat{X}$ we now let $k \to \infty$ thus including all excursions away from zero. This way we can define the martingale increment representation (M_t) over all excursions and thus for all t except for the set of singularities, that is $\{t : \bar{x}_{\xi_i}(t) = 0$ for some $i \in \{1, \ldots, n\}\}$. This implies that the processes $\widehat{X}(t)$ must satisfy the specified martingale problem between singularities.

(c) Completion of the proof of Theorem 1. It remains to complete the proof of the uniqueness of the regular conditional distributions defined by (2.25). In the previous section we proved that $P(\cdot|\bar{X})$ satisfies the martingale problem (2.29), (2.30) for almost all $\bar{X}$, where $\bar{X}$ is uniquely characterized by the the $(\bar{G}, \delta_{\bar{X}})$-martingale problem. This gives a local characterization in intervals of local regularity. From here it is straightforward Ito calculus to see that in fact the $(\widehat{G}_X, \delta_{\widehat{X}})$-martingale problem must be solved (which is a local property).

In order to obtain a global characterization of the conditional distributions it now suffices to verify that the laws of local functionals $\mathcal{H}$ of $\widehat{X}$ at the terminal time T (not involving components where singularities occur at time T) are uniquely determined by this martingale problem. However this was proved in Proposition 1.1, which expresses the moments of $\widehat{X}$ in terms of the dual process $\{\zeta_t\}$. Finally, by a standard argument the uniqueness of the solution to the martingale problem at fixed times yields uniqueness to the law of the full process $\widehat{X}$.

In all these arguments we have assumed that $c_\xi(\cdot) \equiv 0$. By a modification of the standard Girsanov formula for measure-valued processes the well-posedness of the martingale problem when $c_\xi(\cdot) \equiv 0$ for ξ off a compact

set follows from the well-posedness in the case $c_\xi(\cdot) \equiv 0$. The proof of this follows using Theorem 7.2.2 ([D, 93]) and then following the argument of Lemma 10.1.2.1 in [D, 93]. The well-posedness in the general case then follows by an approximation argument by first putting $c_\xi \equiv 0$ off a sequence of finite sets increasing to Ω and then showing that the law in a bounded region is determined by the resulting limit but we will not give the details here. This completes the proof of Theorem 1.
Remark Note that if $t \to \tau$ where $\bar{x}_\xi$ is singular at τ, $\langle M_t \rangle \to \infty$ and the process $\widehat{X}_\xi(t)$ does not converge a.s. However using Lemma 1.8 we can show that it does converge in distribution to a local equilibrium law that depends only on $\{\widehat{X}_\eta(\tau) : \eta \in \partial I\}$.

3. The longtime behavior (Proof of Theorem 2)

The strategy of the proof is to first show part (b) of Theorem 2, which is the ergodic theorem for $\mathcal{L}(X(t))(t \to \infty)$, which will then in a rather standard fashion give also part (a) on invariant measures. Some additional considerations allow to prove part (c) on the total mass and relative weight processes. The part (d) requires then again a different viewpoint.

We focus next on the tools needed for the proof of Theorem 2, part (b). Note first that since it it well-known that for the case where $\hat{a}$ is recurrent $\mathcal{L}(\bar{X}(t))$ converges to $\delta_{\underline{0}}$, this case is trivial and therefore we can focus on the proof of the result for the case where $\widehat{a}$ is *transient*.

In the transient case the main task is, to compare the laws of two processes, which start in different initial states and to show, that they become equal in the (weak) limit $t \to \infty$, if they have initially the same finite intensity measure $\theta \in \mathcal{M}([0,1])$. The proofs of the ergodic theorem for $\mathcal{L}(X(t))$ as $t \to \infty$ will need quite some preparation for the following reason. In the case of multidimensional components the analysis of the longtime behavior in the case of transient migration has in the literature only been carried out for processes having a dual process, such as interacting Fleming-Viot processes. The conditional duality of subsection 1(b) *can not* be used here, since the underlying migration mechanism is not accessible to a straightforward analysis as far as its longtime behavior goes. In systems with one-dimensional components the alternative would be to use coupling. Those techniques are however only working, at least in their classical form, for the one-dimensional components and faces here some difficulties, due to the *multidimensionality* and *noncompactness* of components. Therefore we have to introduce a *new type of coupling dynamics*, which allows to handle *multidimensional* components.

Outline To achieve the goal of comparing different initial states we set up the right framework in subsection 3(a) below and then we proceed in three further main steps. In subsection 3(b) we introduce a notion of distance for multitype populations and based on this concept we construct with quite some technical effort a new type of *coupling* between the two considered processes. In subsection 3(c) we prove the main result of this chapter, namely Proposition 3.1 which says that the coupling is *successful* under appropriate conditions. In subsection 3(d) we complete the Proof of Theorem 2 (a) and (b) based on the coupling results.

(a) The key objects for the proof by coupling. We outline the strategy of the coupling proof now in more detail and introduce the required concepts. The point is to construct a coupling of two processes $X^1(t)$ and $X^2(t)$ with the same dynamics, but starting in different initial states, by defining both processes for all $t \geq 0$ on one common probability space. The idea is to achieve this for multi-dimensional components, by defining the *bivariate process* $(X^1(t), X^2(t))$ as a *functional* of a new process $Z(t)$ with a new type space. This process $(Z(t))_{t \geq 0}$ has the form:

$$Z(t) = (z_\xi(t))_{\xi \in \Omega}, \tag{3.1}$$

with

$$z_\xi(t) \in \mathcal{M}(([0,1] \cup \{*\})^2), \tag{3.2}$$

the latter satisfying that if we define the following functionals of $Z(t)$, called $Z^i(t),\ i=1,2$:

$$
\begin{aligned}
z^1_\xi(t)(\cdot) &= z_\xi(t)(\cdot \times \{[0,1]\cup\{*\}\}) \quad \text{restricted to } [0,1] \\
z^2_\xi(t)(\cdot) &= z_\xi(t)(\{[0,1]\cup\{*\}\}\times \cdot) \quad \text{restricted to } [0,1],
\end{aligned} \tag{3.3}
$$

then $\{z^1_\xi(t)\}_{\xi\in\Omega}, \{z^2_\xi(t)\}_{\xi\in\Omega}$ are versions of the processes $\mathcal{L}(X^1(t))$ respectively $\mathcal{L}(X^2(t))$.

In this set-up the new element $*$ will allow us to encode the total mass difference of the two processes and still work in the enlarged state space with *equal masses* in the two components. We shall define the dynamics of $(Z(t))_{t\geq 0}$ via a martingale problem in Definition 3.1 later on, after preparing the ingredients in step 1-7 of subsection 3(b).

Suppose we are given such a coupled dynamics satisfying (3.1) - (3.3), then we are ready to return to our goal to compare two processes $X^1(t), X^2(t)$, which start in two different initial states but now both are defined on one probability space and both following the dynamics of a branching system with state dependent branching rate $h(\bar{x})$. In order to carry this out we define below a suitable distance function $\|\cdot,\cdot\|$ (cf. 3.6)) and consider then the expected distance:

$$
\Delta_\xi(t) = E\|z^1_\xi(t), z^2_\xi(t)\|. \tag{3.4}
$$

The final goal will be to show that if $\widehat{a}(\xi,\eta)$ is transient and $\mathcal{L}(X^1(0)), \mathcal{L}(X^2(0))$ are translation invariant, spatially ergodic states, with the same finite *atomic* intensity $\theta \in \mathcal{M}_a([0,1])$, then the coupling is successful, meaning:

$$
\Delta_\xi(t) \underset{t\to\infty}{\longrightarrow} 0. \tag{3.5}
$$

With this result available, we then tackle the question of what happens if θ is a non atomic measure by a rather standard approximation argument. The result on (3.5) will be formulated and proved in subsection 3(c), while in subsection 3(b) we face the complicated task to define the dynamics of $Z(t)$.

(b) Construction of a coupling. The construction proceeds in seven steps, in step 1 we define appropriate notions of coupling of two measures and distance between two measures, in step 2 we explain the basic ideas and problems on a particle level, which leads in steps 3 and 4 to the construction of the coupling for the particle model, steps 5 and 6 carry out the diffusion limit of small mass - many particles - rapid branching separately for the migration and the branching. Finally step 7 combines both mechanisms and characterizes the coupled dynamics for the measure valued diffusion via a martingale problem.

Step 1. Matching two measure-valued states In order to compare two states X_1 and X_2 of our system, we need to define a suitable distance function based on a matching of the types. Since we shall later consider elements X_1, X_2, which are random with a translation invariant distribution, we will focus on the notion of distance for the components, i.e. elements of $\mathcal{M}([0,1])$. Here the main point is, that two elements can differ both in the relative weights they assign to various types in $[0,1]$ and in the total mass. We need a notion of difference combining these two effects suitably.

Let $x_1, x_2 \in \mathcal{M}([0,1])$. In order to take into account, that $x_1([0,1])$ and $x_2([0,1])$ may differ we introduce an extra type called $\{*\}$. Let $I^* = [0,1]\cup\{*\}$. Denote by π_i projections on the i-th component of a measure on a product space, with $|_A$ restriction of a measure to the set A. Define:

$$
\|x_1,x_2\| = \inf\left\{ \int\limits_{I^*}\int\limits_{I^*} d(u,v)z(du,dv) \;\middle|\; z\in\mathcal{M}((I^*)^2),\ \pi_1 z\Big|_{[0,1]} = x_1,\ \pi_2 z\Big|_{[0,1]} = x_2 \right\},
$$

with

$$
d(u,v) = \begin{cases} 3 & u\neq v \quad \text{and } u=\{*\} \text{ or } v=\{*\} \text{ but not both,} \\ 1 & u\neq v; \quad u,v\in[0,1] \\ 0 & u=v\in[0,1] \\ \infty & u=v=*. \end{cases} \tag{3.6}
$$

This notion of distance is most useful, if the involved measures are atomic.

Remark The *minimizers* in (3.6) above will be useful for us to construct the coupled dynamics later on. Such a minimizer will assign exactly total mass $|x_1([0,1]) - x_2([0,1])|$ to the points with one component being $\{*\}$.

Remark The distance function just introduced can be compared with the variation norm:

(i) $\|x_1 - x_2\|_{\mathrm{Var}} \le \|x_1, x_2\|$
(ii) If $x_1([0,1]) = x_2([0,1])$ then $\|x_1 - x_2\|_{\mathrm{Var}} = \|x_1, x_2\|$.

In order to carry out above procedure of coupling and association in a stochastic dynamic, both for the particle and the diffusion limit, we have to define, which of the particles in the two processes arising via (3.3) should be matched. This means we have to associate with every pair $x_1, x_2 \in \mathcal{M}([0,1])$ of initial configurations of our two processes a bivariate initial state $\Psi(x_1, x_2) = z \in \mathcal{M}((I^*)^2)$ in an optimal way in the sense of (3.6) to get a good bivariate dynamic. It is important to do this for both the particle model, that is for integer-valued atomic measures, and the diffusion limit, that is for general measures. We now work out this point.

Since we are interested in longtime behavior, in view of Lemma 0.1 there is no loss in generality in assuming that $X(0)$ is pure atomic and we make this assumption throughout the remainder of this section. In other words, we assume that θ is atomic and hence is supported by *countably many types*, which then means that for all times the process $X(t)$ has a support concentrated on these types, so that we deal only with atomic measures $(\mathcal{M}_a([0,1]))$. We will introduce now a coupling of two atomic measures on $[0,1]$, which solves the minimization problem (3.6). The key point is to match particles which do not have the same type in such a way all types are treated in an *exchangeable* fashion. This is formalized as follows.

Let $x, y \in \mathcal{M}_a([0,1])$. Define the map Ψ:

$$\Psi : \mathcal{M}_a([0,1]) \times \mathcal{M}_a([0,1]) \to \mathcal{M}_a(([0,1] \cup \{*\})^2), \tag{3.7}$$

by setting: (we write for the relative weights $\widehat{x}$ of x also $(x)\widehat{\ } = x/\bar{x}$)

$$\begin{aligned} \Psi(x,y)((u,v)) &= (x \wedge y)(u), \qquad \text{for } u = v,\ u \in [0,1] \\ \Psi(x,y)((*,v)) &= (\bar{x} - \bar{y})^-((x-y)^-)\widehat{\ }(v), \quad \text{for } v \in [0,1] \\ \Psi(x,y)((u,*)) &= (\bar{x} - \bar{y})^+((x-y)^+)\widehat{\ }(u), \quad \text{for } u \in [0,1] \\ \Psi(x,y)((u,v)) &= \frac{1}{\overline{(x - x \wedge y)} \vee \overline{(y - x \wedge y)}} [((x-y)^+) \otimes ((x-y)^-)\,]\,(u,v), \\ &\qquad\qquad \text{for } (u,v) \in [0,1]^2,\ u \ne v. \end{aligned} \tag{3.8}$$

Lemma 3.1. *For $x, y \in \mathcal{M}_a([0,1])$, the measure $\Psi(x,y)$ solves the minimization problem (3.6).* □

Proof One explicitly checks that the marginals are concentrated on $[0,1]$ (nothing is put on $(*,*)$) and equal to x respectively y. As little mass as possible is put on coordinates involving * and as much as possible on the diagonal. Then the assertion follows from the construction and the definition of $d(\cdot,\cdot)$.

The crucial observation is that for atomic measures z we have:

Lemma 3.2. *The two properties below characterize minimizers z of (3.6) for atomic measures for given marginals x_1, x_2:*

$$\text{If } u \in [0,1], \text{ then } z((u,w)) \cdot z((w',u)) = 0 \quad \forall w, w' \ne u, \tag{3.9}$$

$$z(\{*\} \times [0,1]) = 0 \quad \text{or} \quad z([0,1] \times \{*\}) = 0. \tag{3.10}$$

The map Ψ chooses among all minimizers of (3.6) for given marginals x_1, x_2, a minimizer with a symmetry (exchangeability) property of the types. □.

Proof If one of the conditions is violated we can put more mass on $[0,1] \times [0,1]$ or the diagonal and hence decrease the integral of $d(\cdot,\cdot)$. For example, in case (3.9) if $z(u,w) > z(w',u) = \delta > 0$ we define

$$\begin{aligned} \tilde{z}(u,u) &= z(u,u) + \delta \\ \tilde{z}(w',u) &= 0 \\ \tilde{z}(u,w) &= z(u,w) - \delta \\ \tilde{z}(w',w) &= z(w',w) + \delta. \end{aligned}$$

Then the mass on the diagonal is increased but both marginals remain unchanged. Therefore the distance is decreased for $\tilde{z}$ which shows that z is not a minimizer. On the other hand one verifies that the minimum value of the integral is attained.

One problem arises if we start with *integer valued* measure in $\mathcal{M}_a([0,1])$ and if we want the minimizer of (3.6) also to be *integer valued*. This is relevant, if we discuss particle versions of our process, which will be necessary to show the existence of a solution to the martingale problem characterizing the process $(Z(t))_{t \geq 0})$, which we want to construct in this subsection. Notice that for x, y integer valued $x \wedge y$ is on the diagonal already integer valued and that the weights off the diagonal of $[0,1] \times [0,1]$ and of the sets $\{*\} \times [0,1], [0,1] \times \{*\}$ of the solution in (3.8) are integer valued. These are exactly the sets where $d(\cdot,\cdot)$ of (3.6) is constant. Therefore we get a nice solution to the minimization problem in the integer valued measures by *randomizing*. Namely consider the following random mechanism independently of everything else.

Consider first the individuals to be matched with $*$. If $\bar{x} > \bar{y} (\bar{x} < \bar{y})$ we draw the second (first) component at random from those not matched on the diagonal. The remaining one we match by drawing one from each component at random and matching them, continue till all are matched.

This leads to a *random map* $\widehat{\Psi}$:

$$\widehat{\Psi}: \quad (\mathcal{M}([0,1]))^2 \to \mathcal{M}(([0,1] \times \{*\})^2) \tag{3.11}$$

such that

$$E\widehat{\Psi} = \Psi. \tag{3.12}$$

Step 2. Coupled branching particle systems - intuitive ideas. We describe and solve the problem of constructing a coupling first for the analogous *particle model* described in the introduction and in this step 2 we roughly sketch the problem and the idea of the solution in the context of the particle model and step 3 and step 4 give the formal construction for particle systems. The execution of this idea in the diffusion limit is somewhat involved and is carried out in steps 5 - 7. Actually the existence proof of the coupled dynamics is based on the particle system approximation of the diffusion.

The main problem in passing in coupling arguments to the *multitype* situation is the fact, that for any coupled dynamics with the structure described in (3.1) - (3.3), the branching rate for type u at a fixed site ξ, can be different in the two component processes even if $x^1_\xi(t)(u) = x^2_\xi(t)(u)$. This is due to the fact that $|h(\bar{x}^1_\xi(t)) - h(\bar{x}^2_\xi(t))|$ depends on all types present at site ξ and hence might still be positive. The only way out of this problem is therefore to guarantee that the evolution mechanism automatically produces a coupling of the total masses, such that this unpleasant rate goes to 0 as $t \to \infty$ (at least in mean). This is achieved as follows.

In order to be able to always work for every component $\xi \in \Omega$, with equal particle numbers in the X^1 and X^2-process, we introduce the ghost particles, which have type * and are only used for the bookkeeping of the excess and deficits of particles in the process X^1 compared to X^2 in the various components $\xi \in \Omega$. Now we should couple the evolutions of X^1 and X^2 to obtain the bivariate dynamics Z.

Consider first the branching, which takes place in the single spatial components and then the migration, which takes place in pairs of spatial components.

Particles of the same type should branch simultaneously in both components $X^1(t), X^2(t)$ at the *maximal possible rate*. The remaining branching is again done to have as much simultaneous branching as possible in both components even if this means simultaneous branching of particles of different types in the two components.

The same principle is applied to the migration, where we have to consider pairs of components $(\xi,\eta)\in\Omega\times\Omega$. In particular we need to try with every jump of a particle pair from ξ to η involving a yet free (i.e. a particle coupled with a ghost particle) particle at site ξ say in process X^1, to first couple the incoming one at site η with the same type (of the other process X^2) and if not possible as second choice with a particle of process X^2 of arbitrary (but different from *) type at site η and only as last choice leave the incoming particle as free particle at the site η that is only matched with a ghost particle.

Let us summarize, the crucial point in this construction is, to associate also elements of different types for the purpose of coupling branching and migration mechanism. For this association of particles of different type during the evolution we need a canonical procedure, which we construct in step 3.

A coupled mechanism as described above, will guarantee that the total masses have the property that $|\bar{x}^1_\xi - \bar{x}^2_\xi|$ does not change due to branching, if $\bar{x}^1_\xi = \bar{x}^2_\xi$, even though the masses attached to types i.e. $(x^1_\xi - x^2_\xi)$ may still change. On the other hand the free particles of X^1 annihilate upon meeting those of X^2 and vice versa, so that the *expected discrepancy in total mass*, that is $E|\bar{x}^1(t)-\bar{x}^2(t)|$, decreases. Hence it makes sense to try to prove, that $|h(\bar{x}^1(t))-h(\bar{x}^2(t))|$ converges to 0 in expectation as time proceeds. Then as a consequence more and more *discrepancies in type* between the processes disappear via migration and dynamical recombination before too many new discrepancies are created. This is a crucial step in generalizing the well-known coupling for one-type situations to the multitype situation. We next begin the formal construction.

Step 3. Coupled particle system, part 1: Dynamical rematching This is a preparatory step. We will construct in this step one of the three key mechanisms of a bivariate evolution for a particle model, which leads to configurations we can describe (in a reduced description) by integer valued measures on $([0,1]\cup\{*\})^2$. This prepares us together with step 4, where the two other mechanisms are introduced, for step 5 and step 6 where we carry out the diffusion limit to show the existence of the process $Z(t)$ of (3.1), which will be characterized by a martingale problem. The present construction of the particle process $(Z_a(t))_{t\geq 0}$ is also more intuitive and serves as an introduction to the technically more demanding step 5.

We will begin in this step 3 by associating masses of the two component processes during the evolution in order to achieve for *each time* that the state is a *minimizer* of (3.6). The definition of the dynamics on $(\mathcal{M}([0,1]\cup\{*\})^2)^\Omega$ based on the rematching mechanism occurring after each branching or migration transition happens in step 4.

In order to define the association we will need as a new ingredient, a *randomized matching map*, which we call $\widehat{\chi}$ but the point is that it needs to be defined only on *small perturbations of minimizers* of (3.6), which arise from a minimizer via a branching or a migration transition of the process in one (respectively two) component. The rematching will take place componentwise. We shall see that this means we deal only with two transitions

$$z \to z + \delta_{(u,v)},\ z \to z - \delta_{(u,v)}, \tag{3.13}$$

where z is a *minimizer*. Observe that a particle leaving or dying keeps a minimizer a minimizer. We define therefore

$$\widehat{\chi}(z-\delta_{(u,v)}) = z-\delta_{(u,v)}, \text{if } z \text{ was a minimizer.} \tag{3.14}$$

Hence we have to deal only with *adding* one new particle, but this already is a bit tricky.

Let z be an element in $\mathcal{M}([0,1]\cup\{*\})^2)$, which is a minimizer of (3.6). We will need to find a new minimizer of this variational problem after adding *one new particle* of type (u,v). In order to find this new random state $\widehat{\chi}(z+\delta_{(u,v)})$, we define the random map

$$\widehat{\chi}: z+\delta_{(u,v)} \to \widehat{\chi}(z+\delta_{(u,v)}), \tag{3.15}$$

by considering separately the different possible configurations.

Consider the following recombinations resulting from using the new particle pair of type (u,v) and one old one taken from the population given by z, to pass to a better matching of the four particles, if this is possible (here $u,v,w \in [0,1]$ and $u \neq v$):

Type of new pair	*Type of old pair*	*new matching*
(u,u)	-	(u,u)
$(u,*)$	$(*,v)$	(u,v)
$(*,v)$	$(v,*)$	(v,v)
(u,v)	(v,u)	$(u,u)+(v,v)$
$(u,*)$	(w,u)	$(u,u)+(w,*)$
$(*,v)$	(v,w)	$(v,v)+(*,w)$
(u,v)	(v,w)	$(v,v)+(u,w)$
(u,v)	(w,u)	$(u,u)+(w,v)$
$(u,*)$	-	$(u,*)$.

This table defines $\widehat{\chi}(z+\delta_{(u,v)})$ as follows. If the new added particle pair with type (u,v) is given we look up the lines which apply to that given type and then we try to realize the matchings, which occur first in the above order of the table read from the top. Among possibilities on the same level we choose the candidates for a new matching among the old pairs of particles, which are allowed on that level, *at random* according to the *uniform distribution.*

Step 4: Coupled particle system, part 2: branching, migration and rematching combined
We are now ready to describe the *evolution of the (bivariate) particle system* started in a state corresponding to a measure z which is a minimizer of (3.6). Before we define the bivariate Markov process of $\mathcal{M}((I^*)^2)$ in terms of its generator, we define a dynamics of the individual particles, which then induces the desired Markov evolution if we simply observe the numbers of particles of the various types.

The definition of the dynamics on $(\mathcal{M}([0,1]\cup\{*\})^2)^\Omega$ happens in three pieces (α) - (γ): first within colonies to handle the *branching* mechanism, second we shall turn in (β) to the *migration*, where we have to consider two different colonies and finally we combine both in (γ).

If in the description below we choose a randomized matching function $\widehat{\chi}$, we choose the randomization at each step independent and independent of everything else.

(α) Start with transitions related to the *branching* mechanism at a single site ξ. Without migration we have no interaction and hence it suffices to define a *one site* dynamics leading to a configuration, which we can describe by an element of $\mathcal{M}((I^*)^2)$.

- Every (u,v) particle performs binary branching at the rate $h((z^1([0,1]))\wedge h((z^2([0,1])))$, where z^1, z^2 are the projections of z on the first resp. second component.
- Every (u,v) particle makes a transition at the rate $|h(z^1([0,1]))-h(z^2([0,1]))|$ and this net transition is created in two steps (taking place at the same time) as follows.

  Consider first the case where $h(\bar{z}_1)\geq(\leq)h(\bar{z}_2)$ if $\bar{z}_1\geq(\leq)\bar{z}_2$.

  (a) Assume first $u,v\neq\{*\}$ and distinguish according to the sign of $\bar{z}^1-\bar{z}^2$.

  In case of the $+$ sign of $\bar{z}^1-\bar{z}^2$ the rule is:

  A $(u,*)$ is created if a birth occurs. If a death occurs a $(*,v)$ particle is created and at the same time (u,v) is removed. Birth and death occur each with probability $\frac{1}{2}$. Call the resulting state z'.
  Then the system switches in the second step to the state given by $\widehat{\chi}(z')$. The state $\widehat{\chi}(z')$ looks as follows.

  In the case of a birth, $\widehat{\chi}(z')$ results in a new $(u,*)$, in case that either $u\neq v$ or there exists no $w\neq u$ such that the pair (w,u) is present. If on the other hand $u=v$ and there exists $w\neq u$ with the pair (w,u) present then we get by rematching $+(u,u)-(w,u)+(w,*)$. Here w is chosen at random according to $\hat{z}((I\backslash\{u\})\times\{u\})$.

  In case of death in the loss of one $(u,*)$, if there was one (which then combines with $(*,v)$ to compensate the loss of $u,v)$). If there was none, we pick a $(w,*)$ if available, which is combined with

$(*,v)$ to form a new (w,v). Here w is chosen according to $\widehat{z}(\cdot,*)$ normalized to a probability measure. The net result is $(w,v)-(u,v)$. If no w is available the result is $(*,v)-(u,v)$.

In case of the – sign of $\bar{z}^1-\bar{z}^2$ the rule is:
The role of the components 1 and 2 in the above are simply exchanged.
(b) In the case where $(u,v)=(u,*)$ or $(*,v)$ we simply have birth or death of $(u,*)$ $((*,v))$ at rate $(h(z^1([0,1]))-h(z^2([0,1])))^+$ respectively $((\dots)^-)$ with probability $\frac{1}{2}$ for birth and death. A rematching is not possible if z was a minimizer.

In the case of $h(\bar{z}_1)\geq(\leq)h(\bar{z}_2)$ if $\bar{z}_2\geq(\leq)z_1$, the modification is only needed for the situation described in (a) above and is simple in case of the + sign: a $(*,v)$ is created if a birth occurs and a $(u,*)$ once a death occurs and in the latter case (u,v) is removed. Otherwise proceed now as described above.

The above dynamics induces a one site coupled multitype branching particle system. This process is well-defined. If we now pass to the reduced description by specifying at time t only the number of particles of the various types in $I^*\times I^*$, we obtain a Markov process $(z_a(t))_{t\geq 0}$ on $\mathcal{M}((I^*)^2)$. It is important later on that this process satisfies a martingale problem, which we now describe.

We will write the generator in a particular form, that will allow us later on to handle the diffusion limit. This requires us to consider separately those transitions which involve a type $(u,*)$ or $(*,u)$ in the event where the current weight on this type is 0. This way the generator has two pieces ($\widetilde{G}^a_{bra}$ and $\widetilde{H}$), one which in the diffusion limit will lead to a differential operator and one involving an increasing process associated to a local time functional. This decomposition is needed in the limit, since otherwise passing to the diffusion limit we cannot preserve the positivity of the states, as will become clear in step 5.

First we define the generator $\widetilde{G}^a_{bra}$ by:

$$\widetilde{G}^a_{bra}=\widetilde{G}^{a,inc}_{bra}+\widetilde{G}^{a,dec}_{bra}, \tag{3.16}$$

The two components are given as follows. Define:

$$\begin{aligned}
&\chi^{ij}_{h,I}(z)=\\
&[h(z^i([0,1]))-h(z^j([0,1]))]1(z^i([0,1])>z^j([0,1]))1(h(z^i([0,1]))>h(z^j([0,1]))),\ \ i\neq j\in\{1,2\}\\
&\chi^{ij}_{h,D}(z)=\\
&[h(z^j([0,1]))-h(z^i([0,1]))]1(z^i([0,1])>z^j([0,1]))1(h(z^j([0,1]))<h(z^j([0,1]))),\ \ i\neq j\in\{1,2\}.
\end{aligned} \tag{3.17}$$

Then define for F twice continuously differentiable,

$$\begin{aligned}
\widetilde{G}^{a,inc}_{bra}(F)(z)\ \ =&\\
&1/2\int\limits_{I^*\times I^*}[F(z+\delta_{(u_1,u_2)})+F(z-\delta_{(u_1,u_2)})\\
&\qquad -2F(z)](h(z^1(I))\wedge h(z^2(I)))z(d(u_1,u_2))\\
&+\chi^{12}_{h,I}(z)\cdot\int\limits_{I\times I^*}[F(z+\delta_{(u,*)})+F(z-\delta_{(u,*)})-2F(z)]1(z(\{u,*\})>0))z(d(u,v))\\
&+\chi^{21}_{h,I}(z)\cdot\int\limits_{I^*\times I}[F(z+\delta_{(*,u)})+F(z-\delta_{(*,u)})\\
&\qquad -2F(z)]1(z(\{*,u\})>0))z(d(v,u)).
\end{aligned} \tag{3.18}$$

Similarly define $\widetilde{G}^{a,dec}_{bra}$ by replacing $\chi^{i,j}_{h,I}$ by $\chi^{ij}_{h,D}$ but incorporate the following minor modification of the above. In this case extra births of minority type lead to a decrease in the $*$ type and extra deaths lead to an increase of the $*$-type.

Finally define the generator involving the transitions of types involving $*$ in a configuration with the weights $z((*,u))=0$ or $z((u,*))=0$ (compare second and third term in (3.18)) by:

$$\begin{aligned}\widetilde{H}^I_F(z) \quad = \quad & +\chi^{12}_{h,I}(z)\cdot\{\int\limits_{I\times I^*}[F(z+\delta_{(u,*)})-F(z)]1(z(\{(u,*)=0\}))z(d(u,v)) \\ & \quad +\int\limits_I\int\limits_{I\times I}[F(z-\delta_{(w,*)}-\delta_{(u,v)}+\delta_{(w,v)})-F(z)]1(z(\{(u,*)=0\}))z(d(u,v))\hat{z}(d(w,*))]\} \\ & +\chi^{21}_{h,I}(z)\cdot\{\int\limits_{I\times I}[F(z+\delta_{(*,u)})-F(z)]1(z(\{(*,u)=0\}))z(d(v,u)) \\ & \quad +\int\limits_I\int\limits_{I\times I}[F(z-\delta_{(*,w)}-\delta_{(v,u)}+\delta_{(v,w)})-F(z)]1(z(\{(*,u)=0\}))z(d(v,u))\hat{z}(d(*,w))]\}.\end{aligned}$$

The case of $\widetilde{H}^D_F$ is the analog to the above. We abbreviate by $\widetilde{H}_F$ the sum $\widetilde{H}^I_F+\widetilde{H}^D_F$ and we define the following functional of the process $(z_a(t))_{t\geq 0}$:

$$\widetilde{H}_F(z_a(t),t)=\int\limits_0^t\widetilde{H}_F(z_a(s))ds. \tag{3.19}$$

Then the process $(z_a(t))_{t\geq 0}$ satisfies the following martingale problem. Let F be a bounded function F, which is twice continuously differentiable with compact support on $\mathcal{M}((I^*)^2)$. Then:

$$M^a_F(t):=F(z_a(t))-F(z_a(0))-\int_0^t\widetilde{G}^a_{bra}F(z_a(s))ds-\widetilde{H}_F(z_a(t),t) \tag{3.20}$$

is a martingale.

Since we also need the interacting case we have to consider the evolution of all components $(z_\xi(t))_{\xi\in\Omega}$ each independently following the above one site dynamics. We define the process generator G^a_{bra} by setting for F a bounded function on $\mathcal{M}((([0,1]\cup\{*\})^2)^\Omega)$, which depends only on finitely many components:

$$G^a_{bra}(F)(Z)=\sum_{\xi\in\Omega}\widetilde{G}^a_{bra}(F\circ\pi_\xi)(Z),\quad H_F(Z(t),t)=\sum_{\xi\in\Omega}\widetilde{H}_{F\circ\pi_\xi}(z_\xi(t),t), \tag{3.21}$$

where $F\circ\pi_\xi:\mathcal{M}(([0,1]\cup\{*\})^2)\to\mathbb{R}$ is obtained by viewing F as a function of z_ξ.

(β) Next we come to the *migration* part. Again we specify first the motion of individual particles. The *migration* of a single particle from ξ to η is carried out by carrying out instantaneously (independently for every particle) the following three transitions with transition rate $\bar{a}(\xi,\eta)$ (recall $\bar{a}(\xi,\eta)=a(\eta,\xi)$).

- A particle with type (u,v) moves from ξ to η.
- Once the particle of type (u,v) arrives at the new colony η, which was in an optimal state z, an instantaneous transition to $\widehat{\chi}(z+\delta_{(u,v)})$ occurs with $\widehat{\chi}$ as defined in step 3.
- At the colony ξ the new state is equal to $\tilde{z}-\delta_{(u,v)}$ if $\tilde{z}$ was the old one and was a minimizer.

This defines a unique particle jump process on Ω, with particles of different types and with migration rate independent of the type.

If we consider only the number of particles of the various types in $I^*\times I^*$ one finds at time t in a colony ξ, we obtain a Markov process on $\mathcal{M}(([0,1]\cup\{*\})^2)$. We describe next this Markov process by its generator.

For this purpose we consider transitions resulting from a particle of type (u,v) which migrates from site η to site ξ chosen such that the *states at all times satisfy the minimization conditions* (recall Lemma 3.2). Note that such a transition occurs at rate $z_\eta((u,v))a(\eta,\xi)$. The resulting contribution to the change of the configuration due to the state at ξ is obtained according to (expectation w.r.t. the randomization in $\widehat{\chi}$)

$$E(\widehat{\chi}(z_\xi+\delta_{(u,v)})-z_\xi). \tag{3.22}$$

We write this expectation in the form (decomposing according to whether the type w or both types w and w' are involved in the rematching taking place after the migration transition):

$$\sum_w \lambda_\xi((u,v),w)K_\xi((u,v),w)\ 1(C_\xi) \qquad \sum_{w'}\sum_w \lambda_\xi((u,v),(w,w'))K_\xi((u,v),(w,w'))1(C_\xi), \tag{3.23}$$

where the signed measure λ_ξ describes the effect of the transition at ξ if the rematching involves the types w or w and w' and K_ξ specifies the (conditional) probability of this rematching transition to be chosen among all possible ones and finally $C_\xi = C_\xi((u,v),w)$, respectively $C_\xi = C_\xi((u,v),(w,w'))$, describe the respective sets of configurations of particles in which the indicated reaction is possible.

We now specify these quantities λ_ξ and K_ξ depending on C_ξ. We distinguish the cases

$$\begin{array}{lll} (a) & (u,v) & \text{lies on the diagonal or in } ([0,1]\times\{*\})\cup(\{*\}\times[0,1]) \text{ and} \\ (b) & (u,v) & \text{lies in the remaining part of } I^*\times I^*. \end{array} \tag{3.24}$$

Case (a) Then transitions at site ξ to a new state as indicated in the table below occur in a row only if z_ξ does not satisfy the conditions in the higher rows but does satisfy the condition in the given row, that is, $1(C_\xi((u,v),w))) = 0$ if the condition in any of the higher rows is satisfied (minimization property!).

Imm.	New state at ξ -old at ξ	Conditional Prob.	Condition
	λ_ξ	K_ξ	C_ξ
(u,u)	$\delta_{(u,u)}$	1	none
$(u,*)$	$-\delta_{(*,u)}+\delta_{(u,u)}$	1	$z_\xi(*,u)\neq 0$
$(u,*)$	$-\delta_{(*,w)}+\delta_{(u,w)}$	$z_\xi(*,w)/z_\xi(*,[0,1]\backslash\{u\})$	$z_\xi(*,[0,1]\backslash\{u\})\neq 0$
$(u,*)$	$-\delta_{(w,u)}+\delta_{(u,u)}+\delta_{(w,*)}$	$z_\xi(w,u)/z_\xi([0,1]\backslash\{u\},u)$	$z_\xi([0,1]\backslash\{u\},u)\neq 0$
$(u,*)$	$\delta_{(u,*)}$	1	other
$(*,u)$	$-\delta_{(u,*)}+\delta_{(u,u)}$	1	$z_\xi(u,*)\neq 0$
$(*,u)$	$-\delta_{(w,*)}+\delta_{(w,u)}$	$z_\xi(w,*)/z_\xi([0,1]\backslash\{u\},*)$	$z_\xi([0,1]\backslash\{u\},*)\neq 0$
$(*,u)$	$-\delta_{(u,w)}+\delta_{(u,u)}+\delta_{(*,w)}$	$z_\xi(u,w)/z_\xi(u,[0,1]\backslash\{u\})$	$z_\xi(u,[0,1]\backslash\{u\})\neq 0$
$(*,u)$	$\delta_{(*,u)}$	1	other

Case (b) The following table is implemented according to the same rules as the one above except in rows 4 and 5 where we randomize which to realize if they are both applicable:

Imm.	New state at ξ - old at ξ	Conditional Prob.	Condition
	λ_ξ	K_ξ	C_ξ
(u,v)	$+\delta_{(u,u)}+\delta_{(v,v)}+\delta_{(w,w)}$ $-\delta_{(u,v)}-\delta_{(w,u)}-\delta_{(v,w)}$	$z_\xi(v,w)z_\xi(w,u)/Z$ Z norm factor	$\bigcup_w\{z_\xi((w,u))>0, z_\xi((v,w))>0\}$
(u,v)	$\delta_{(u,u)}+\delta_{(v,v)}-\delta_{(v,u)}$	1	$z_\xi(v,u)\neq 0$
(u,v)	$+\ \delta_{(u,u)}+\delta_{(v,v)}+\delta_{(w,w')}$ $-\delta_{(u,v)}-\delta_{(w,u)}-\delta_{(v,w')}$	$z_\xi(v,w)z_\xi(w'u)/Z$ Z norm factor	$z_\xi((v,[0,1]))>0$, $z_\xi(([0,1],u))>0$, not previous row
(u,v)	$\delta_{(u,u)}+\delta_{(w,v)}-\delta_{(w,u)}$	$z_\xi(w,u)/z_\xi([0,1]\backslash\{u\},u)$	$z_\xi([0,1]\backslash\{u\},u)\neq 0$
(u,v)	$\delta_{(v,v)}+\delta_{(u,w)}-\delta_{(v,w)}$	$z_\xi(v,w)/z_\xi(v,[0,1]\backslash\{v\})$	$z_\xi(v,[0,1]\backslash\{v\})\neq 0$
(u,v)	$\delta_{(u,v)}$	1	other

In all cases the state at η has the simultaneous transition $z_\eta \to z_\eta - \delta_{(u,v)}$ at rate $z_\eta(u,v)$ for the appropriate (u,v).

What we have now achieved is that we can define the migration part of the generator for the bivariate evolution started in minimal states. Namely the following transitions occur (averaged over the randomisation) at rates:

$$\begin{aligned} Z &\longrightarrow Z+\lambda_\xi((u,v),w)-\delta_{\eta,(u,v)}, &&\bar a(\eta,\xi)K_\xi((u,v),w)z_\eta((u,v))1\!\!1(C_\xi((u,v),w))) \\ Z &\longrightarrow Z+\lambda_\xi((u,v),(w,w'))-\delta_{\eta,(u,v)}, &&\bar a(\eta,\xi)K_\xi((u,v),(w,w'))z_\eta((u,v))1\!\!1(C_\xi((u,v),(w,w'))) \end{aligned} \tag{3.25}$$

By summing over $((u,v),w),(u,v),(w,w'))$ and then both η and ξ we obtain the total of contributions to the generator, which we denote by

$$G^a_{mig} \tag{3.26}$$

and which acts on functions F on $\mathcal{E}^* \subseteq (\mathcal{M}((I^*)^2))^\Omega$ which are depending only on finitely many components.

(γ) Now we are ready to define the Markov process $(Z_a(t))_{t\geq 0}$, which arises by letting the coupled branching and the coupled migration dynamic, which we defined sofar, occur independently. This defines uniquely a stochastic process $(Z_a(t))_{t\geq 0}$ which is Markov and stays within the minimal states. The following property of this process will be important. Define on bounded functions F depending on finitely many components the operator

$$G^a_{coup} = G^a_{bra} + G^a_{mig}. \tag{3.27}$$

Then $(Z_a(t))_{t\geq 0}$ solves the *martingale problem*:

$$(F(Z_a(t)) - F(Z_a(0)) - \int_0^t G^a_{coup}(F)(Z_a(s))ds - H_F(Z_a(t),t))_{t\geq 0} \tag{3.28}$$

is a martingale for all bounded F which depend on finitely many of the coordinates.

The next point is now to pass from the particle model $(Z_a(t))_{t\geq 0}$ to the *diffusion limit*, $(Z(t))_{t\geq 0}$, and to characterize the limiting law by a martingale problem. The process $Z(t) = \{z_\xi(t) : \xi \in \Omega\}$ will be constructed in step 7 via a martingale problem, whose formulation requires the following objects: a first order *differential operator* denoted G^d_{mig} (introduced in step 5), a second order differential operator G^d_{bra} and *local time functionals* denoted $(L(t))_{t\geq 0}$ (introduced in step 6). The local time functionals come from the coupled branching at those sites where the weight of $(u,*)$ or $(*,v)$ is 0 and appears in the expressions for $H_F(Z_a(t),t)$. Since the proof of the convergence of the particle system to a diffusion system that satisfies the limiting martingale problem is carried out following the standard arguments, we will sketch some parts the proofs but focus on the special features involved. First, since the coupling mechanisms introduce somewhat complicated expressions, we look separately at the effects of branching and migration. In the next step we focus on ingredients for G^d_{mig} and the definition of G^d_{mig} and in the subsequent step on $(G^d_{bra}, L(t))_{t\geq 0})$.

Step 5. The diffusion limit: The coupled pure migration dynamics In this step we consider the diffusion limit of the migration part of the dynamics of Z_a alone. This means taking the diffusion limit of the process described by (3.28) with $h \equiv 0$, i.e. no branching. The diffusion limit arises by giving every particle mass ε and increasing the initial number of particles to $\{\lfloor \varepsilon^{-1} z_\xi \rfloor,\ \xi \in \Omega\}$. The increased number of particles means that the number of migration transitions between two sites diverges so that a law of large numbers applies leading to the following mass flow described by a first order infinitesimal generator. Namely the limiting (first order) differential generator for the process of masses is given by

$$G^d_{mig}F(z) = \sum_\xi \sum_\eta \bar{a}(\eta,\xi) \int_{I^*\times I^*} z_\eta(d(u,v)) \left\{ -\frac{\partial F(z)}{\partial z_\eta}(u,v) + \int_{I^*\times I^*} \Lambda(z_\xi,(u,v),d(u',v')) \frac{\partial F(z)}{\partial z_\xi}(u',v') \right\}, \tag{3.29}$$

where the measures $\Lambda(z_\xi,(u,v),\cdot)$ are given by

$$\begin{aligned} \Lambda(z_\xi,(u,v),d(u',v')) = &\left[\int_I \lambda_\xi((u,v),w) 1(C_\xi((u,v),w)) K_\xi((u,v),dw) \right] d(u',v') \\ &+ \left[\int_I \int_I \lambda_\xi((u,v),(w,w'))((u,v)(w,w')) K_\xi((u,v)(d(w,w')) 1(C_\xi) \right]. \end{aligned} \tag{3.30}$$

and the expression in square brackets is the signed measure defined in (3.23) and the subsequent table.

The operator G^d_{mig} acts on functions F on $(\mathcal{M}(\{[0,1] \cup \{*\}\}^2))^\Omega$ which depend on finitely many components $\xi \in \Omega$ and are C^2 as functions of those measure valued components (recall (0.45)). In fact the system of coupled differential equations describing the flow has a unique solution if we start the system in a coupling of two configurations in ζ. This is standard theory of differential equations with values in Banach spaces. We leave the details to the reader.

We can check the tightness of the process $(Z^a_\varepsilon(t))_{t \geq 0}$ in the case of vanishing branching rate ($h \equiv 0$) as elements of $D([0,\infty), \mathcal{E})$ and then we can prove by applying the law of large numbers to the numbers of migration steps of a particle from ξ to η and with denoting with $(Y^d(t))_{t\geq 0}$ the deterministic flow defined by the first order differential operator G^d_{mig}, that for vanishing branching rate

$$\mathcal{L}\left((Z^a_\varepsilon(t))_{t\geq 0}\right) \underset{\varepsilon \to 0}{\Longrightarrow} \mathcal{L}\left((Y^d(t))_{t\geq 0}\right). \tag{3.31}$$

To check the tightness we simply observe first that a single transition involving the state at a fixed site is bounded by ε since particles leave or move in. Furthermore the expected transition rate at ξ in the particle system which started in the state in $\mathcal{E}$ is bounded by $C \cdot \varepsilon^{-1}$ with C independent of ε since the expected number of particles is bounded (in time) in finite time intervals. Therefore we know that in any finite window or observation the expected local jump rates are bounded uniformly by Const ε^{-1}. As a consequence the expected local total variation of the total mass configuration over bounded time interval $[0,T]$ is bounded by $D \cdot T$ and D a constant independent of ε and therefore the quadratic variation goes to 0 as $\varepsilon \to 0$. This implies the tightness by standard criteria.

Step 6 The diffusion limit: The coupled pure branching mechanism. We next consider the diffusion limit of the coupled branching mechanism at a *fixed site* starting from the particle martingale problem (3.20). Consider the process $(z_a(t))_{t\geq 0}$ and assign to each particle mass ε, speed up the branching rate per particle by ε^{-1} and take as initial distribution $[z_a(0)/\varepsilon]$ with $[\cdot]$ Gauß bracket. The limiting process $(z(t))_{t\geq 0}$ will be a diffusion, which lives on the nonnegative measures and which will be characterized by a martingale problem involving two main objects namely (1) a generator (a second order differential operator) and (2) local time functionals of certain weights at 0 in order to keep the states nonnegative in types $(u,*)$ and $(*,v)$.

We define first the second order operator. We use the $\sim$ to indicate that we focus on a *single* site only (no interaction). Let F be a twice continuously differentiable function on $\mathcal{M}([0,1])$. Set

$$\begin{aligned}
\widetilde{G}^d_{bra} F(z) &= \widetilde{G}^{d,inc}_{bra} + \widetilde{G}^{d,dec}_{bra} \\
\widetilde{G}^{d,inc}_{bra} F(z) &= \Bigg\{ \frac{1}{2} \int\limits_{I^2 \times I^2} \left[\frac{\partial^2 F(z)}{\partial z \partial z}((u_1,u_2)(v_1,v_2)) \right] (h(z^1(I)) \wedge h(z^2(I)) \\
&\quad \delta_{(u_1,u_2)}(d(v_1,v_2)) z(d(u_1,u_2)) \\
&\quad + \chi^{12}_{h,I}(z) \cdot \int\limits_{(I^*)^2 \times I} \left[\frac{\partial^2 F(z)}{\partial z \partial z}((u,*),(v_1,v_2)) \right] (h(z^1(I)) - h(z^2(I)))^+ 1(z(\{u,*) > 0\}) \\
&\quad \delta_{(u,*)}(d(v_1,v_2)) z(d(u,v)) \\
&\quad + \chi^{21}_{h,I} \cdot \left[\int\limits_{(I^*)^2 \times I} \frac{\partial^2 F(z)}{\partial z \partial z}((*,u),(v_1,v_2)) \right] (h(z^1(I)) - h(z^2(I)))^- 1(z(\{*,u) > 0\}) \\
&\quad \delta_{(*,u)}(d(v_1,v_2)) z(d(v,u)) \Bigg\}
\end{aligned}$$

and $\widetilde{G}^{d,dec}_{bra}$ is obtained by replacing $\chi^{i,j}_{h,I}$ by $\chi^{i,j}_{h,D}$.

Acting on functions F with bounded second derivatives, the approximating and limiting generators differ by an error term that goes uniformly to zero as $\varepsilon \to 0$. Also note that although the generators involve indicator functions, the mappings $z \to \widetilde{G}^d_{\text{bra}} F(z)$ are continuous (because the associated weight factors go to zero as the boundary is approached). The key subtle point is the situation in which $z^1(I) > z^2(I)$ but $z(*,u) = 0$ (or $z^1(I) < z^2(I)$ and $z(*,v) = 0$. The idea here is to introduce a reflection mechanism so that $z(u,*) = 0$ stays nonnegative but at the same time $z^1(I) - z^2(I)$ is non-increasing. To do this we introduce increasing processes and the associated *local time functionals*, $\ell^1(t)(u,*), \ell^2(t)(*,v)$ with $u,v \in [0,1]$, the collection of which will be abbreviated $L(t)$.

We also need the approximating objects on the particle level which are given by:

$$\ell_a^{1,\varepsilon}(u,*)(t) := \int_0^t 1(\{z_a^\varepsilon(s)((u,*))=0\})ds, \quad \ell_a^{2,\varepsilon}(*,v) = \int_0^t 1(\{z_a^\varepsilon(s)((*,v))=0\})ds \tag{3.32}$$

and $L^\varepsilon(t) = \{\ell^{1,\varepsilon}(u,*)(t), \ell^{2,\varepsilon}(*,v)(t) | u,v \in [0,1]\}$.

Remark The role of the local time functionals is similar to that in Skorokhod's construction of reflecting Brownian motion. The process $Z(t)$ should always have non-negative weight in every point of $([0,1] \cup \{*\})^2$. On the other hand the first term in the expression for $\widetilde{G}^d_{bra}$ produces positive diffusion constants in $(*,u)$ or $(u,*)$, even if the weight of $z(t)$ in these points is 0 and would lead therefore without compensation to negative values. In order to understand this phenomenon return to step 4 and note that as far as $h(\bar{z}^1) \neq h(\bar{z}^2),\ \bar{z}^2 \leq \bar{z}^1$, we see for example from the particle construction in step 4 that the $(*,\cdot)$-types have (before rematching takes place) positive increasing process but due to the rematching the all-over result is a nonnegative state (formal calculations follow below). Therefore to keep $z(t)$ nonnegative, we need to introduce some reflection mechanism for each type in each colony for states of the form $(u,*)$ or $(*,v)$. This reflection is achieved using local times and corresponds on the particle level to the rematching procedure described in step 4.

We will show next that $(z_a^\varepsilon(t), L^\varepsilon(t)) \underset{\varepsilon\to 0}{\Longrightarrow} (z(t), L(t))$ as processes and we will show that $((z(t),L(t)))_{t\geq 0}$ satisfies a martingale problem. We write $\ell^1(u,*), \ell^2(*,u)$ with $u \in I$ for the components of L (recall (3.32)). Note that $\ell^1(u,*)(t)$ is a nondecreasing function of t and we view it as a measure in the time variable.

The above remark shows that the following functional of $z(t)$ is important, which replaces in the continuum limit the functional $\widetilde{H}_F(z_a(t),t)$, namely:

$$\begin{aligned}
\widetilde{H}_F(z(t),t) := \quad & \Bigg[\int_0^t \chi^{12}_{h,I}(z(s)) \int_I \tfrac{\partial F(z(s))}{\partial z}((u,*)) 1(z(s)((u,*))=0) z(s)(d(u,*)) \ell^1(u,*)(ds) \\
& + \int_0^t \int_{I^2\times I^*} \Big\{ -\tfrac{\partial F(z(s))}{\partial z}((w,*)) - \tfrac{\partial F(z(s))}{\partial z}((u,v)) + \tfrac{\partial F(z(s))}{\partial z}((w,v)) \Big\} \\
& \qquad 1(z(s)((u,*))=0) z(s)(d(u,v)) \hat{z}(s)(d(w,*)) \ell^1(u,*)(ds) \Bigg] \\
+ \quad & \Bigg[\int_0^t \chi^{21}_{h,I}(z(s)) \int_I \tfrac{\partial F(z(s))}{\partial z}((*,u)) 1(z(s)((*,u))=0) z(s)(d(*,u)) \ell^2(*,u)(ds) \\
& + \int_{I^2\times I^*} \Big\{ -\tfrac{\partial F(z(s))}{\partial z}((*,w)) - \tfrac{\partial F(z(s))}{\partial z}((v,u)) + \tfrac{\partial F(z(s))}{\partial z}((v,w)) \Big\} \\
& \qquad 1(z(s)((*,u))=0) z(s)(d(v,u)) \hat{z}(s)(d(*,w)) \ell^2(*,u)(ds) \Bigg] \\
& + \Bigg[\int_0^t \chi^{12}_{h,D}(z(s)) \int_I \tfrac{\partial F(z(s))}{\partial z}((u,*)) 1(z(s)((u,*))=0) z(s)(d(u,*)) \ell^1(u,*)(ds) \\
& + \int_{I^2\times I^*} \Big\{ -\tfrac{\partial F(z(s))}{\partial z}((w,*)) + \tfrac{\partial F(z(s))}{\partial z}((w,v)) \Big\} \\
& \qquad 1(z(s)((u,*))=0) z(s)(d(u,v)) \hat{z}(s)(d(w,*)) \ell^1(u,*)(ds) \Bigg] \\
& + \Bigg[\int_0^t \chi^{21}_{h,D}(z(s)) \int_{I\times I^*} \tfrac{\partial F(z(s))}{\partial z}(*,u) 1(z(s)((*,u))=0) z(s)(d(*,u)) \ell^2(*,u)(ds) \\
& + \int_{I^2\times I^*} \Big\{ -\tfrac{\partial F(z(s))}{\partial z}((*,w)) + \tfrac{\partial F(z(s))}{\partial z}((v,w)) \Big\} \\
& \qquad 1(z(s)((*,u))=0) z(s)(d(v,u)) \hat{z}(s)(d(w,*)) \ell^2(*,u)(ds) \Bigg].
\end{aligned} \tag{3.33}$$

What we shall need now is that for a suitable sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ tending to zero, we obtain a limiting process satisfying a martingale problem so that we can perform calculations. In fact more is true:

Lemma 3.3. *The limit of $(z_a^\varepsilon(t), L^\varepsilon(t))_{t\geq 0}$ for $\varepsilon \to 0$ exists. The limit has the form*

$$(z(t), L(t)) = \Big(\{z(t)(u,v)) : (u,v) \in I^* \times I^*\}, \quad \{\ell_a^1((u,*))(t)), \ell_a^2(*,v)(t), u,v \in I^* \times I^*\}\Big) \tag{3.34}$$

and satisfies the following martingale problem:

For every twice continuously differentiable function F on $\mathcal{M}((I^*)^2)$ define:

$$M_F(z(t),t) := F(z(t)) - F(z(0)) - \int_0^t \widetilde{G}^d_{bra}(F)(z(s))ds - \widetilde{H}_F(z(t),t). \tag{3.35}$$

Then $(M_F(z(t),t))_{t\geq 0}$ is a continuous martingale, where:

$$z(0) = z_0, \quad \text{and } z(t) \geq 0 \qquad \forall t \geq 0, \tag{3.36}$$

$$\left(\ell^1((u,*))(t)\right)_{t\geq 0}, \left(\ell^2((*,v))(t)\right)_{t\geq 0} \quad \text{are increasing processes for } u,v \in [0,1] \quad \text{with} \tag{3.37}$$

$$\int_0^t z(s)((u,*))\ell^1((u,*))(ds) = 0 \quad \forall\, t \geq 0,\ u \in [0,1]. \tag{3.38}$$

and

$$\int_0^t z(s)((*,v))\ell^2((*,v))(ds) = 0 \quad \forall\, t \geq 0,\ v \in [0,1]. \qquad \square \tag{3.39}$$

Remark It is general fact, that the description of Z of the above definition implies, that $\ell^1(u,*)(t)$ is the local time of $z(t)$ in $(u,*)$ till time t, in the sense that the occupation time formula holds. For a semimartingale $(Y_t)_{t\geq 0}$ this formula says $\int_0^t \Phi(Y_s)d\langle Y,Y\rangle_s = \int_{-\infty}^{+\infty} \Phi(a)L_t(a)da$, where $L_t(b)$ is the local time in b up to time t and Φ a continuous function.

Proof of Lemma 3.3 Tightness and convergence of the branching part follows from standard arguments as are used showing the convergence of branching processes to branching diffusions. We leave out these well-known details. Similarly the convergence to the limiting expression $\widetilde{G}^d_{bra}$ is standard. The next point is the term related to $L(t)$ and to $\widetilde{H}_F$.

We first turn to the local times themselves and note that for every $u \in [0,1]$ we have to show tightness of the process given by the expression $(\ell_a^{1,\varepsilon}((u,*))(t))_{t\geq 0}$ and for every fixed time t we have to show that in distribution

$$\int_0^t 1(\{z_a^\varepsilon(s)(\{u\}\times\{*\}) = 0)ds \underset{\varepsilon\to 0}{\Longrightarrow} \ell^1((u,*))(t). \tag{3.40}$$

This is proved by standard arguments. (In particular, it is analogous to the well-known case of the convergence of random walk local time to Brownian local time. See [RY], Chapt. VI, Exer. 2.11 and [R], Theorem 10.1.)

If we have the convergence of the local time functional the next point is the actual convergence of the term corresponding to $\widetilde{H}_F$. Again we need tightness of the sequence of functionals and then convergence. We first focus on the aspect of verifying that the local time properties and expressions appearing in the limiting martingale problem fit together, that is, we get convergence.

First of all the properties (3.37), (3.38),(3.39) are true in the particle process z_a. Namely the first two are obvious from the construction and the last two follow then from the defining relation (3.40). Finally to convince ourselves that the terms in $\widetilde{H}_F$ are correct we proceed as follows.

Recall that we consider the case in which the z_ξ are all *atomic* measures on $[0,1]$. In order to see how the reflection at the value 0 in points of the form $(*,u)$ or $(v,*)$ in the limiting martingale problem is achieved, we return to the particle model. We translate the transitions in the particle model which occur if $z((w,*))=0$ or $z((*,v))=0$ due to a birth or death at (w,v) in just one of the two components into the corresponding changes in the diffusion limit.

Consider the transitions, which can occur in the event that at a fixed time t, $z((w,*))=0$ and $\bar{z}^2 \leq \bar{z}^1$ or $z((*,v))=0$ together with $\bar{z}^2 \geq \bar{z}^1$. In this constellation the evolution of the particle model involves various types at once due to the instantaneous improvement of the matching. The above two cases therefor split into two further cases depending on the relation between $h(\bar{z}^1), h(\bar{z}^2)$.

Consider the configuration:

$$z((w,*))=0, \quad \bar{z}^2 \leq \bar{z}^1, \quad h(\bar{z}^2) > h(\bar{z}^1). \tag{3.41}$$

Then at the type (w,v) a birth or death at the second component may occur and leads after matching again to the following change of the state z_ξ (abbreviate by $\hat{z}^1((\cdot,*)) = z((\cdot,*))/z(I\times\{*\})$:

$$\begin{array}{lll} \text{birth} & +(u,v)-(u,*) & u \text{ chosen with frequency } \widehat{z}^1((u,*)) \\ \text{death} & +(w,*)-(w,v). & \end{array} \tag{3.42}$$

Next consider the case

$$z((w,*))=0, \quad \bar{z}^2 \leq \bar{z}^1, \quad h(\bar{z}^2) < h(\bar{z}^1). \tag{3.43}$$

A change at the type (w,v) a birth or death might occur in the first component and leads after matching to

$$\begin{array}{lll} \text{birth} & +(w,*) & \\ \text{death} & +(u,v)-(u,*)-(w,v) & u \text{ chosen at rate } \widehat{z}^1((u,*)). \end{array} \tag{3.44}$$

Altogether we always get after averaging over the probability for birth or death the result

$$(u,v)+(w,*)-(u,*)-(w,v). \tag{3.45}$$

It is the term $(w,*)$, which will on the event $\bar{z}^2 \leq \bar{z}^1$ be absorbed in the local time functional, so that the following term remains which gives rise to three different first order (drift) terms which arise in the functional H_F in the diffusion limit:

$$(u,v)-(u,*)-(w,v) \quad u \text{ chosen with frequency } \widehat{z}^1((u,*)). \tag{3.46}$$

Note that all the changes involve choosing also the v, accompanying the fixed type w in the first process. This means that the changes in the first case (resp. second case) have to be picked with the following frequencies:

$$\widehat{z}^2((w,v)),\ (\widehat{z}^1((u,w))). \tag{3.47}$$

This yields the terms $\widetilde{H}_F$ in the particle model and now we have to verify the convergence.

¿From the convergence in (3.40) and law of large number effects arising since all the randomizations occurring in the rematching take place a diverging number of times, we can conclude the convergence of the expressions involving $\widetilde{H}_F$ to those involving H_F. This argument is in substance not different from that in the classical case of constructing reflected Brownian motion using a martingale problem involving local time functionals. We leave the details to the reader.

Finally we have to make a similar calculation for the configuration with $z((*,v))=0$ and the event $\bar{z}^2 \geq \bar{z}^1$, we omit the details, which are now straightforward from the above. This completes the proof of Lemma 3.3.

Step 7 Diffusion limit: The coupled system with both branching and migration. Now we have defined all the ingredients needed to define the bivariate process $Z(t)=\{z_\xi(t)\}_{\xi\in\Omega}$ on the *type space* $([0,1]\cup\{*\})^2$ and with the required marginal processes (see (3.3)) and *coupled dynamics*. We lift $\widetilde{G}^d_{bra}, \widetilde{H}_F$ to

G^d_{bra}, H_F according to the same recipe as in the particle system in (3.21). We next define (recall (3.29) and (3.32)):

$$G^d_{coup} = G^d_{bra} + G^d_{mig}. \tag{3.48}$$

and what we mean by the martingale problem induced by G_{coup} and $(L_t)_{t \geq 0}$. Then using the previous two steps we show that a solution exists.

Definition 3.1. Consider $L = (L(t))_{t\geq 0} = (\{\ell^{\xi,1}(\bullet)(t), \ell^{\xi,2}(\bullet)(t); \xi \in \Omega\})_{t \geq 0}$ where for $i = 1, 2$ and fixed t, the $\ell^{\xi,i}(\bullet)$ are functions on $[0,1] \times \{*\}$ and $\{*\} \times [0,1]$. We define a law P on $C(([0,\infty),(\mathcal{M}(J))^\Omega)$ to be the solution of the $((G_{coup}, L), Z_0)$-martingale problem if the following holds:

$$Z(0) = Z_0, \tag{3.49}$$

$$Z(t) \geq 0 \quad \text{a.s}, \tag{3.50}$$

$$t \rightarrow L(t) \quad \text{is increasing}, \tag{3.51}$$

$$\int\limits_0^t z_\xi(s)((u,*))\ell^{\xi,1}((u,*))(ds) = 0 \quad \forall u \in [0,1],\ t \geq 0,\ \xi \in \Omega, \tag{3.52}$$

$$\int\limits_0^t z_\xi(s)((*,v))\ell^{\xi,2}((*,v))(ds) = 0 \quad \forall v \in [0,1],\ t \geq 0,\ \xi \in \Omega,$$

and the process below is a martingale for all F depending twice continuously differentiable on finitely many components:

$$\left\{F(Z(t)) - F(Z_0) - \int\limits_0^t (G_{coup}F)(Z(s))ds - H_F(Z(t),t)\right\}_{t \geq 0}. \qquad \square \tag{3.53}$$

We need to verify that these definitions make sense and to establish that there exists a solution to the $((G_{coup}, L), \delta_Z)$-martingale problem which provides the required coupling. Note however that we do *not* need uniqueness of the martingale problem at this point.

Lemma 3.4.
(a) *Given two deterministic initial states X^1, X^2 in $\mathcal{E}$, there exists on $C[[0,\infty), (\mathcal{M}(([0,1] \cup \{*\})^2))^\Omega]$ a solution P_{Z_0} of the $((G_{coup}, L), \delta_{Z_0})$-martingale problem with $Z_0 = \{\Psi(x^1_\xi, x^2_\xi),\ \xi \in \Omega\}$, where Ψ is as in (3.8).*
(b) *If $z_\xi(0)$ is a minimizer of (3.6) $\forall \xi$, then:*

$$z_\xi(t) \quad \text{is a minimizer of (3.6)}\ \forall t \geq 0,\ \xi \in \Omega. \tag{3.54}$$

(c) *The functional $(Z^1(t), Z^2(t))$ of the process $(Z(t))_{t\geq 0}$ as defined in (3.3) has components which are versions of the two systems of branching with interaction defined in (0.18), provided that the projections of $Z(0)$ on the two components are given by $X^1(0)$ and $X^2(0)$ respectively.* $\square$

Proof
(a) The proof proceeds by showing that a solution of the martingale problem arises as diffusion limit, of the particle system $(Z_a(t))_{t\geq 0}$ defined in step 4. The essential steps have already been indicated in step 5 and in Lemma 3.3. In particular we consider the particle model defined in step 4 and then we associate with every particle mass ε, speedup the branching by ε^{-1} and use as initial states a sequence of initial configurations given as follows. Set $Z_0 = \{[z_\xi/\varepsilon], \xi \in \Omega\}$, where [ ] is the Gauss bracket, and use this as initial configuration for the particle dynamics. This way, for every $\varepsilon > 0$, we obtain a particle processes $(Z^{(\varepsilon)}_a(t))_{t \geq 0}$.

In order to ensure the tightness of this family in the space $D([0,\infty), \mathcal{E}^2)$ (where $\mathcal{E}^2 \subseteq (\mathcal{M}((I^*)^2))^\Omega$ is defined by requiring that the projections on both I^*-components are in $\mathcal{E}$) follows from the fact that the projections of $Z^\varepsilon(t)$ have the property that their norm (defining $\mathcal{E}$) is bounded uniformly in ε and in times $t \leq T$. Since we know the tightness of the single component processes for dynamic from step 5 and step 6 we can lift these results to $\mathcal{E}^2$-valued processes.

On the one hand we showed in the previous steps that without interactions, i.e. for $a(\xi, \eta) \equiv \delta(\xi, \eta)$, we obtain convergence of the branching system in law as $\varepsilon \to 0$ to a solution to the limiting martingale problem and on the other hand the migration part alone converges as $\varepsilon \to 0$ to limiting dynamics given by a Feller-semigroup. Combining these arguments we can show that limit points of the probability laws of the approximating particle systems satisfy the limiting martingale problem associated with $G^d_{coup}F + \ H_F + G^d_{mig}F$.

(b) We return to the particle model and note that the evolution satisfies for all $t > 0$ the relations and (3.9), (3.10) by its very construction, if this is true for $t = 0$. Hence it stays in the set of minimizers during the evolution. Obviously, any limit point of the probability laws of the particle systems inherits this property, so that the state is almost surely a minimizer of (3.6).

(c) In order to see that for any limit point the laws of the marginal processes satisfy the martingale problem (0.18) observe that in the the first, resp. second, components of the approximating particle systems have the required migration rates and branching at rate $h(\bar{\eta}^1_\xi)$ resp. $h(\bar{\eta}^2_\xi)$ by construction. Again this property is inherited by any limit point.

(c) The coupling is successful. Next we come to the main goal of this subsection 3. Recall that the coupled process $(Z^1(t), Z^2(t))_{t \geq 0}$ is a functional of the process $(Z(t))_{t \geq 0}$, which we defined in (3.49) - (3.53), where the functional is given by (3.3). Our main result says that we have constructed a *successful coupling*. Precisely:

Proposition 3.1. *Assume that $\mathcal{L}(X^1(0))$ and $\mathcal{L}(X^2(0))$ are both translation invariant and linear ergodic (see 3.82) with intensity measure $\theta \in \mathcal{M}([0,1])$. Then if $\widehat{a}(\xi, \eta)$ is transient the following holds.*
(a) If in addition $\theta \in \mathcal{M}_a([0,1])$ then with Δ_ξ as in (3.4):

$$\lim_{t \to \infty} \Delta_\xi(t) = 0. \tag{3.55}$$

(b) If θ is not atomic then still

$$\mathcal{L}(X^1(t)) - \mathcal{L}(X^2(t)) \underset{t \to \infty}{\Longrightarrow} 0 - \textit{measure}. \qquad \square \tag{3.56}$$

Proof of Proposition 3.1 The strategy is to prove first (a) and then use an approximation argument to obtain (b).

For the proof of the Proposition part (a) we need first some preparation and we formulate the needed facts in Lemma 3.5 and Lemma 3.6, which we both prove before we start the actual proof of the Proposition 3.1 in steps 2 and 3. Let us mention that in the sequel in all arguments and differential equations we get the same type of expressions if we work with the approximating particle process $(Z^\varepsilon_a(t))_{t \geq 0}$. In fact this is the easiest way to derive some of the differential equations.

Step 1: Preparations It is convenient to first reformulate things a bit. In order to show that $\Delta_\xi(t) \to 0$ as $t \to \infty$, it will turn out below that it suffices to show that the total masses and the relative weights of types approach each other. In formulas this looks as follows. Define:

$$k(t) = E|\bar{z}^1_\xi(t) - \bar{z}^2_\xi(t)| \tag{3.57}$$

$$K(t) = E[z_\xi(t)(\{(u,v) | u, v \in [0,1] \cup \{*\}, \ u \neq v\})]. \tag{3.58}$$

Then we shall see later that, (3.55) is equivalent to showing (recall that $(*,*)$ has weight 0 under $z_\xi(t)$ a.s.):

$$K(t) \underset{t \to \infty}{\longrightarrow} 0. \tag{3.59}$$

The first task needed to do this will be to analyze the behavior of $k(t)$, which is the key to controlling $E|h(\bar{z}^1_\xi(t) - h(\bar{z}^2_\xi(t)|$, which creates in the coupled dynamics the jumps away from the diagonal.

For the process of total masses we can use a result by Cox and Greven ([CG1, CG2]) to prove after preparing some tools below at the very end of this step 1:

Lemma 3.5. *Assume that $\widehat{a}$ is transient. If $\mathcal{L}(X^1(0))$ and $\mathcal{L}(X^2(0))$ are translation invariant and ergodic and if they agree in the intensity of their total mass, then*

$$k(t) \searrow 0 \text{ as } t \to \infty. \qquad \square \tag{3.60}$$

The second task is to move from the total masses to the distribution of mass onto the various types. Based on this definition of $Z(t)$ and its generator consisting of the operator G_{coup} and the local times $(L(t))_{t\geq 0}$ (see (3.53)), we shall now derive a differential equation for the quantities $k(t)$ and $K(t)$ (recall (3.57) (3.58) for the definition), which will then in the sequel allow the analyses of the behavior of $K(t)$ for $t \to \infty$.

For the purpose of writing down the differential equations we need various quantities, first for the total mass process, then for the process itself. Define for the total mass functional (where sign(t) is $+1$ for $t > 0$, -1 for $t < 0$ and 0 for $t = 0$):

$$G(t) := -2\sum_{\eta} a(\xi,\eta) E\left[|\bar{z}^1_\eta(t) - \bar{z}^2_\eta(t) \mid 1\!\mathrm{I}\ \left(\text{sign}\ \left[\bar{z}^1_\eta(t) - \bar{z}^2_\eta(t)\right] \neq\ \text{sign}\ \left[\bar{z}^1_\xi(t) - \bar{z}^2_\xi(t)\right]\right)\right], \tag{3.61}$$

which describes the mean rates at which mass excesses and deficiencies annihilate each other and which does not depend on ξ because of the translation invariance of the law of $Z(t)$. For the process $Z(t)$ itself define the following two expected effects of unequal branching rates in the components aggregated over types on the diagonal of $[0,1]^2$ abbreviated by D and on $\{*\} \times [0,1] \cup [0,1] \times \{*\}$ (the second arises in states in which $\bar{z}^1_\xi \geq \bar{z}^2_\xi$ and h is locally decreasing and hence an increase on the diagonal can occur, or the complementary situation occurs):

$$\begin{aligned} I(t) &= \int\limits_{[0,1]^2} E\left[|h(\bar{z}^1_\xi(t)) - h(\bar{z}^2_\xi(t))| 1\!\mathrm{I}(u=v) z_\xi(t)(\{du,dv\})\right] \\ &= E\left[|h(\bar{z}^1_\xi(t) - h(\bar{z}^2_\xi(t)| z_\xi(t)(D)\right] \end{aligned} \tag{3.62}$$

$$\begin{aligned} J(t) &= \frac{1}{2}\int\limits_{[0,1]} E\Big[\Big\{(h(\bar{z}^1_\xi(t)) - h(\bar{z}^2_\xi(t))^- 1\!\mathrm{I}(z_\xi(t)(\{v,*\}) > 0) \\ &\qquad + (h(\bar{z}^1_\xi(t)) - h(\bar{z}^2_\xi(t))^+ 1\!\mathrm{I}(z_\xi(t)(\{*,u\}) > 0)\Big\} 1\!\mathrm{I}(u \neq v) z_\xi(t)(du,dv)\Big] \\ &\qquad + \frac{1}{2}\int\limits_{[0,1]} E\Big[\Big\{(h(\bar{z}^1_\xi(t) - h(\bar{z}^2_\xi(t))^+ 1\!\mathrm{I}(z_\xi(t)(\{v,*\} > 0) \\ &\qquad + h(\bar{z}^1_\xi(t)) - h(\bar{z}^2_\xi(t)))^- 1\!\mathrm{I}(z_\xi(t)(\{*,u\}) > 0)\Big\} 1\!\mathrm{I}(u \neq v) z_\xi(t)(du,dv)\Big] \end{aligned} \tag{3.63}$$

$$= \frac{1}{2} E\left[|h(\bar{z}^1_\xi(t) - h(\bar{z}^2_\xi(t))| z_\xi(t)([0,1]^2 \backslash D)\right]. \tag{3.64}$$

Furthermore define the mean effect of migration on the matching of types (recall (3.23)):

$$\begin{aligned} h_{\xi,\eta}(t) &= \int\limits_{[0,1]^2} \Big\{2E\big[z_\eta(t)(d(u,v)) 1\!\mathrm{I}(\{u \neq v,\ z_\xi(t)(v,u) > 0\})\big] \\ &\qquad + 2E\big[z_\eta(t)(\{d(u,v) | u \neq v, \exists w \neq u, \bar{w} \neq v : \\ &\qquad\qquad z_\xi(t)(w,u) > 0, \text{ or } z_\xi(t)(v,\bar{w}) > 0)\}, \ z_\xi(t)(\{v,u\}) = 0\}\big]\Big\} \end{aligned} \tag{3.65}$$

$$H(t) := \sum_\eta a(\xi,\eta) h_{\xi,\eta}(t). \tag{3.66}$$

Note that H does not depend on ξ, since the initial distribution (and hence every law for $t > 0$ as well) is translation invariant.

With this notation the following differential equations or inequalities hold:

Lemma 3.6.

$$\frac{d}{dt}k(t) = G(t) \tag{3.67}$$

$$k(t) \to 0 \textit{ as } t \to \infty \quad \textit{implies } I(t) \to 0 \textit{ as } t \to \infty \tag{3.68}$$

$$\frac{d}{dt}K(t) = I(t) - H(t) - J(t) \le I(t). \qquad \square \tag{3.69}$$

Proof of Lemma 3.6

(3.67): The total mass processes $(\bar{z}^1_\xi(t), \bar{z}^2_\xi(t))$ derived from $(Z(t))_{t\ge0}$ by considering $z_\xi([0,1] \times \{[0,1] \cup \{*\}\})$ and $z_\xi(\{[0,1] \cup \{*\}\} \times [0,1])$, follow a certain coupled dynamics for one-dimensional components, which are systems of interacting diffusions. These diffusions arise using a particular diffusion matrix, which follows from the martingale problem (3.53). Namely written as SDE:

$$\begin{aligned} d\bar{z}^1_\xi(t) &= \sum_\eta a(\xi,\eta)(\bar{z}^1_\eta(t) - \bar{z}^1_\xi(t))dt \\ &\quad + (g(\bar{z}^1_\xi(t)) \wedge g(\bar{z}^2_\xi(t))dw^1_\xi(t) + (g(\bar{z}^1_\xi(t)) - g(\bar{z}^2_\xi(t)))^+ dw^2_\xi(t) \end{aligned} \tag{3.70}$$

$$\begin{aligned} d\bar{z}^2_\xi(t) &= \sum_\eta a(\xi,\eta)(\bar{z}^2_\eta(t) - \bar{z}^2_\xi(t))dt \\ &\quad + (g(\bar{z}^1_\xi(t)) \wedge g(\bar{z}^2_\xi(t))dw^1_\xi(t) + (g(\bar{z}^1_\xi(t)) - g(\bar{z}^2_\xi(t)))^- dw^2_\xi(t) \end{aligned} \tag{3.71}$$

where $\{(w^i_\xi(t))_{t\ge0},\ \xi \in \Omega,\ i = 1,2\}$ are i.i.d. Brownian motions. This follows immediately observing that by construction the joint branching induces two coupled diffusions and the branching of only one component happens in disjoint time intervals so that we get the same law if we use the same driving Brownian motions $(w^2_\xi(t))_{t\ge0}$ for both components.

Then the assertion (3.67) follows via two observations. By Ito's calculus and the fact that the local time of $(\bar{z}^1_\xi(t) - \bar{z}^2_\xi(t))$ on the diagonal is 0, since the diffusion constant disappears on that set and is locally Lipschitz as long as the drift for the difference is not identically 0 we obtain a system of coupled differential equations which we can simplify using the translation invariance of the law of the coupled process. See [CG1, 94], Lemma 3.2 for details.

(3.69): This follows from the form of the pre-generator i.e. G_{coup} and $(L(t))_{t\ge0}$. First of all it is easy to see that the migration term results in the terms $H(t)$, due to the translation invariance of the law of the process (see [CG2, 94] Lemma 3 for this calculation). The diffusion term and $L(t)$ contributes with the term $I(t) - J(t)$. Here $I(t)$ reflects the disturbance of the successfully coupled state on the diagonal, which arises in the particle model from a birth or death in only one of the components due to the discrepancy $|h(\bar{x}^1_\xi(t)) - h(\bar{x}^2_\xi(t))|$. On the other hand $J(t)$ reflects some effect in the opposite direction, namely jumps onto the diagonal due to the birth, death in one component which allows a better matching. We omit the straightforward details. This result is alternatively easily verified from using the particle approximation.

(3.68): Since $g(x)/x^2$ converges to 0 as $x \to \infty$ and $g(x)/x$ to $h(0)$ as $x \to 0$, we can bound $h(x) = g(x)/x$ by $C(1 + o(x))$. Therefore, since $E\bar{z}_\xi(t) = \bar{\theta}\ \forall t \ge 0$, we know that $E|h(\bar{z}^1_\xi(t) - h(\bar{z}^2_\xi(t)|$ is uniformly integrable. Furthermore, since $h(x)$ is continuous, $k(t) \underset{t\to\infty}{\longrightarrow} 0$ implies that:

$$E|h(\bar{z}^1_\xi(t)) - h(\bar{z}^2_\xi(t))| \to 0 \quad \text{as } t \to \infty. \tag{3.72}$$

Proof of Lemma 3.5 First recall that for the total masses we get the system ((3.70)- (3.71)) of SDE's describing the total masses of the coupled process. We observe, that for the quantity $E(|z^1_\xi(t) - z^2_\xi(t)|)$ we get according to (3.67) the *same system of differential equations* as in [CG2, 94] Lemma 3.2 and therefore the argument given in this paper carries over completely for (3.60). The boundedness of the components assumed there is irrelevant, since for $\hat{a}$ transient the second moment measures remain bounded and the result carries over (compare Shiga [S]).

Step 2: Proof Proposition 3.1, Part (a)

In the sequel we shall occasionally use second moment measures $E[z_\eta(t)(u)z_\xi(t)(u)]$. It is well-known that for $\hat{a}$ transient this expression is finite if this is the case at $t = 0$. The latter we can assume w.l.o.g. since by monotonicity and comparison with $\{z_\xi(0)(\cdot) \wedge n\}_{\xi\in\Omega}$ we can easily reduce to that case. We omit the standard details and assume in the sequel

$$E\left[(z_\xi(0)(u))^2\right] < \infty. \tag{3.73}$$

Part 1 (liminf) We show first that (recall that here the intensity measure θ is atomic):

$$\liminf_{t\to\infty} \Delta_\xi(t) = 0. \tag{3.74}$$

The proof proceeds indirect. Assume that $\liminf_{t\to\infty} \Delta_\xi(t) = \delta > 0$. Pick a sequence $t_n \uparrow \infty$ such that $\Delta_\xi(t_n) \to \delta$ as $n \to \infty$. Next pick a subsequence $t_{n_k} \uparrow \infty$ such that (note that the $E\bar{z}^i_\xi(t),\ i = 1, 2$ are constant in t and ξ)

$$\mathcal{L}\Big(Z_{t_{n_k}}\Big) \underset{k\to\infty}{\Longrightarrow} \widehat{\nu}. \tag{3.75}$$

Since the second moment measure remains bounded in the case where $\widehat{a}$ is transient, we can conclude with (3.75) and (3.60) that first moments are uniformly integrable in t and hence:

$$E_{\widehat{\nu}}|\bar{z}^1_\xi - \bar{z}^2_\xi| \equiv 0. \tag{3.76}$$

Consider now the quantities $\widehat{I}, \widehat{J}, \widehat{k}, \widehat{H}, \widehat{G}, \widehat{K}$ which are defined as before but now based on the initial law $\widehat{\nu}$ for the coupled dynamics. Then by (3.76) and (3.67) we know that $\widehat{k} \equiv 0$ and hence $\widehat{I} \equiv 0$, and $\widehat{J} \equiv 0$, therefore we have from (3.69) that:

$$\frac{d}{dt}\widehat{K}(t) = -\widehat{H}(t) \le 0. \tag{3.77}$$

By assumption we know that (recall that $\widehat{k} = 0$!) $\widehat{K}(0) = \delta$. Hence either $\widehat{H}(t) \equiv 0$ or $\widehat{K}(t)$ is *strictly* decreasing over some finite time interval and hence eventually smaller than δ. In the latter case we then must have for t large enough and some $\delta' < \delta$:

$$K(t_n + t) < \delta \quad \text{for infinitely many } n, \tag{3.78}$$

which, since $k(t_n + t) \to 0$ as $n \to \infty$, is a contradiction to the assumption that δ is the liminf of $\Delta_\xi(t)$ as $t \to \infty$ and hence is at most $\liminf$ of $\Delta_\xi(t_n + t)$ as $n \to \infty$. Therefore we must have $\widehat{H}(t) \equiv 0$.

Next we show that the relation $\widehat{H} \equiv 0$ implies that z^1 and z^2 are ordered for all types u with $\theta(u) > 0$. First we shall see below, that under the measure $\widehat{\nu}$ the configurations $(z^1_\xi)_{\xi\in\Omega},\ (z^2_\xi)_{\xi\in\Omega}$ must be ordered with respect to any given type u with $\theta(u) > 0$ and with respect to the set $U_\xi = \{\eta \in \Omega | a(\xi, \eta) > 0\}$. This means that with $\widehat{\nu}$-probability one:

$$z^1_\eta(u) \ge z^2_\eta(u) \quad \forall \eta \in U_\xi, \quad \text{or} \quad z^1_\eta(u) \le z^2_\eta(u) \quad \forall \eta \in U_\xi. \tag{3.79}$$

To see this recall (3.66) and (3.9). Then it follows from $\hat{k} \equiv 0$ that either $z^1_\eta(u) = z^2_\eta(u)$ or that for all $t \ge 0$ either $\{(u, v) | v \in [0, 1], v \ne u, v \ne *\}$ or $\{(v, u) | v \in [0, 1], v \ne u, v \ne *\}$ have positive weight under z_η for $\eta \in U_\xi$.

Then the weight of these sets is either the excess or defect of $(z^1_\eta - z^2_\eta)$ w.r.t. type u. For all components connected by $a(\xi,\eta)$ it must be either all excess or all defect, since otherwise $\widehat{H}$ would not be identically zero.

We shall show below after completing the argument that the *irreducibility* of $a(\cdot,\cdot)$ implies that in fact for every ordering $\geq$ with respect to a given type u satisfying $\theta(u) > 0$ (i.e. $z^1_\xi(u) \geq z^2_\xi(u) \quad \forall \xi \in \Omega \Longleftrightarrow z^1 \geq z^2$) one has

$$\widehat{\nu}(\{z^1 \geq z^2 \text{ or } z^1 \leq z^2\}) = 1, \text{ for all } u \text{ with } \theta(u) > 0. \tag{3.80}$$

We shall show next that (3.80) implies that actually equality holds, the reason being that $\widehat{\nu}$ is spatially linear ergodic (we recall this notion in (3.82)). Note that the ergodicity of the initial law implies that, in the case of $\widehat{a}$ transient, for all later times the law is spatially linear ergodic. The latter is in fact only a statement about the original (uncoupled) process. To see this assertion assume first that $E[(\bar{x}_\xi(0))^{2+\varepsilon}] < \infty$. This will imply that $\sup_{t\geq 0} E[\bar{x}_\xi(t)]^{2+\tilde{\varepsilon}} < \infty$ for some $\tilde{\varepsilon} > 0$. This boundedness of $(2+\widetilde{\varepsilon})$-th moment was shown in Lemma 2.2. part (a) in [CGS1, 95]. Then it suffices to show that (for θ atomic)

$$\left\{\sup_{t\geq 0} E[x_\xi(t)(u)x_\eta(t)(u)]\right\} \underset{|\eta-\xi|\to\infty}{\longrightarrow} (\theta(u))^2 \quad \text{and } \liminf_{t\to\infty} E[x_\xi(t)(u)x_\eta(t)(u)] \geq (\theta(u))^2. \tag{3.81}$$

This holds by (3.104) - (3.107) below. Therefore if we can show that $\sup_t E[\bar{x}_\xi(t)]^{2+\varepsilon} \leq \infty$ for some $\varepsilon > 0$ then

$$\sum_\eta a_t(\xi,\eta)x_\eta(u) \underset{t\to\infty}{\longrightarrow} (\theta(u)) \quad \text{in } L_2(\widehat{\nu}_i), \quad i = 1,2, \tag{3.82}$$

if $\widehat{\nu}_i$ denotes the projection on the i-th component.

The assumption on the $(2+\varepsilon)$-th moments for the initial state is finally removed in a standard way by truncation (i.e. passing to the initial state $(z_\xi(0)(u) \wedge n)_{\xi\in\Omega}$) and coupling between the truncated and real process. See [G, 91] for such arguments.

Relation (3.82) and the fact that θ is atomic together with (3.80) imply immediately that

$$\widehat{\nu}(\{z^1 = z^2\}) = 1. \tag{3.83}$$

The relation implies that in fact

$$\widehat{\nu}(\{\widehat{K} = 0\}) = 1. \tag{3.84}$$

But then $K(t_n) \to 0$ as $n \to \infty$, which contradicts the assumption that $\liminf \Delta_\xi(t) = \delta > 0$. Hence (3.74) is proved, provided we verify (3.80)

In order to finally prove (3.80) we fix u and write z_ξ for $z_\xi(\{u\})$. It suffices to show the following: for fixed $T > 0$, which is small enough $\widehat{\nu}(z^1_\xi < z^2_\xi,\ z^1_\zeta = z^2_\zeta,\ z^1_\eta > z^2_\eta) > 0$ for ξ,ζ,η such that $a(\xi,\zeta) > 0,\ a(\zeta,\eta) > 0$, implies there exists a $t_0 \in [0,T]$ such that the system started in $\widehat{\nu}$ satisfies

$$P\left(\left\{z^1_\xi(t_0) < z^2_\xi(t_0),\ z^1_\zeta(t_0) \neq z^2_\zeta(t_0),\ z^1_\eta(t_0) > z^2_\eta(t_0)\right\}\right) > 0. \tag{3.85}$$

But then $\widehat{H}(t)$ would not be identically 0. Therefore $\widehat{\nu}(z^1_\xi > z^2_\xi,\ z^1_\zeta = z^2_\zeta,\ z^1_\xi < z^2_\xi) = 0$ for ξ,ζ,η as described.

Next we can use the irreducibility to get the assertion. Hence the point is to establish (3.85). This is obvious in the particle picture but very tedious in the diffusion limit.

In order to get (3.85) we have to see that the t_0 can be chosen such that conditioned on the event that at time 0 we see the configuration $\{z^1_\xi(u) < z^2_\xi(u),\ z^1_\zeta(u) = z^2_\zeta(u),\ z^1_\eta(u) > z^2_\eta(u)\}$, then with positive probability

$$z^1_\xi(t_0) < z^2_\xi(t_0) \quad \text{and } z^1_\eta(t_0) > z^2_\eta(t_0) \tag{3.86}$$

and furthermore that on the above event the increasing process

$$\langle z^1_\zeta(u)(t) - z^2_\zeta(u)(t)\rangle \quad \text{is not identically 0 in } [0,T]. \tag{3.87}$$

In fact this suffices to hold on the event where the inequalities in (3.86) holds for *all* $t \in [0, t_0]$ with positive probability. This means we have to show that on the event described with positive conditional probability $\langle z^1_\sigma(u)(t) - z^2_\sigma(u) \rangle$ is bounded from above for $\sigma = \xi, \eta$ and from below for $\sigma = \zeta$ and in addition $\sum_\sigma a(\sigma, \sigma')x_\sigma(u)(t)$ is bounded from above for $\sigma = \xi, \eta$. But since we know that conditioning on the event in (3.85) still yields a conditional law supported on distributions in $\mathcal{E}^2$ if this was the case at time 0, this follows by moment calculation. The details of such an argument have been carried out at length in the Appendix of [CG2, 94] we omit further details here.

Part 2 (limsup) We show here that based on step 1

(3.88) $$\limsup_{t\to\infty} \Delta_\xi(t) = 0.$$

Assume that (again we give an indirect argument)

(3.89) $$\limsup_{t\to\infty} \Delta_\xi(t) = \delta > 0.$$

We can then conclude that $limsup_{t\to\infty} K(t) = \delta > 0$. Then from the result of step 1 and the fact that $K(t)$ is smooth (for example a straightforward generator calculation shows that the function K is in $C^2([0,\infty))$) we conclude that there exists a sequence (t_n):

(3.90) $$t_n \text{ is a local maximum of } K(t), \quad K(t_n) \underset{n\to\infty}{\longrightarrow} \delta > 0.$$

Therefore we read of from (3.90) and (3.69) that

(3.91) $$H_{t_n} \leq I_{t_n}.$$

But since $I_{t_n} \to 0$ as $n \to \infty$ by (3.60) and (3.68), we know using $H_t \geq 0$ by definition:

(3.92) $$\lim_{n\to\infty} H_{t_n} = 0.$$

Let $\widehat{\nu}$ be a weak limit point as $n \to \infty$ of the sequence (recall the mean measure is preserved):

(3.93) $$\{\mathcal{L}(Z_{t_n})\}_{n\in\mathbb{N}}.$$

Then denoting again by $\widehat{\ }$ the quantities based on the initial law $\widehat{\nu}$ for the coupled dynamics we get as in the previous step:

(3.94) $$\widehat{H}(0) = 0.$$

As in the argument of the previous step ((3.79) - (3.84)) we conclude that then $\widehat{K}(0) = 0$, which contradicts the assumption that (recall $\widehat{k} = 0$) $\lim_{n\to\infty} \Delta_\xi(t_n) = \delta' > 0$, which requires $\widehat{K}(0) > 0$. Hence (3.88) is proved and this proves part (a) of the Proposition 3.1.

Step 3: Proof of Proposition 3.1, Part (b) The final point is to drop the assumption of θ being atomic. Here we use an approximation by the atomic case. Start by approximating the intensity measure. For $\theta \in \mathcal{M}([0,1])$ we consider a sequence of atomic measures, namely we take the partitions

(3.95) $$\{\frac{j}{2^k} | j = 0, \dots, 2^k\}$$

and define $\theta_{(k)}$, by

(3.96) $$\theta_{(k)} = \sum_{j=1}^{2^k} \left\{ \theta\Big(\Big[\frac{j-1}{2^k},\ \frac{j}{2^k}\Big)\Big) \right\} \delta_{j_k} + \theta(\{1\})\delta_1, \quad j_k = \frac{j}{2^k}.$$

To get a corresponding sequence of approximating processes we replace all the states in $X(0)$ by the same recipe as above. In fact due to the exchangeability of the types under the dynamics and due to the conservation of types (no mutation), we know that

(3.97) $$\mathcal{L}\Big((X(t))_{(k)}\Big) = \mathcal{L}\Big(X(t) | X(0) = X_{(k)}(0)\Big).$$

Furthermore by construction (recall (3.96)) and the fact that types are not changed by the evolution:

$$\mathcal{L}\Big((X(t))_{(k)}\Big) \underset{k\to\infty}{\Longrightarrow} \mathcal{L}(X(t)). \tag{3.98}$$

The latter convergence is uniform in $t \in [0,\infty)$ for every test function defining the weak topology. Therefore combining (3.97) and (3.98) it is easy to obtain with part (a) of Proposition 3.1 the statement, that under the assumption of Proposition 3.1:

$$\mathcal{L}(X^1(t)) - \mathcal{L}(X^2(t)) \underset{t\to\infty}{\Longrightarrow} 0\text{-measure}. \tag{3.99}$$

This completes the proof of Proposition 3.1.

(d) Proof of Theorem 2. (i) Part (a) and (b)

It is best to start with the proof of part (b) of the theorem in both the recurrent and the transient case.

The recurrent case as far as part (b) goes, is trivial due to the extinction result for the total mass process proved in [CFG], so that we focus on the transient case.

The claim of part (b) of the theorem in the transient case follows from Proposition 3.1. Namely note that every translation invariant measure can be represented as a mixture of shift ergodic measures which are the extreme points in that simplex. Therefore it suffices in order to prove part (b) of the theorem to identify the behavior of the shift ergodic components of an initial measure.

Next turn to part (a) of the theorem. Again the recurrent part is simple, it is a consequence of the local extinction result for the total mass process. Consider the case of $\widehat{a}$ transient next. Denote by $\mathcal{T}_\theta$ translation invariant linear ergodic laws with intensity measure θ. By the convergence result there can be in $\mathcal{T}_\theta$ at most one invariant measure for a given finite intensity θ. With the convergence relation of Proposition 3.1 it is then straightforward and standard to construct the equilibrium measure as follows.

We take the initial state $\delta_{\underline{\theta}}$ (every component equals the intensity measure θ) and observe that the set $\{\mathcal{L}(X(t)),\ t \geq 0\}$ is (because of the preservation of the mean intensity measure of a component) relatively compact in the weak topology and hence we can define for a convergent subsequence indexed by $t_n \to \infty$:

$$\nu_\theta = w - \lim_{n\to\infty} \mathcal{L}(X(t_n)). \tag{3.100}$$

In order to show that ν_θ is an invariant measure we show below the following lemma by a second moment calculation.

Lemma 3.7. *Assume that $\widehat{a}$ is transient. If the initial law $\nu \in \mathcal{T}_\theta$ and the total mass in a component has finite second moments, then this holds for every $t > 0$. The distribution ν_θ defined in (3.100) is linear ergodic, that is for every $f \in C([0,1])$ we have under ν_θ:*

$$\sum_{\eta\in\Omega} a_t(\xi,\eta)\langle x_\xi, f\rangle \underset{t\to\infty}{\longrightarrow} \langle\theta, f\rangle \quad \textit{in probability and } L_2. \tag{3.101}$$

The convergence in probability and L_1 holds if we only require ν is ergodic with intensity θ. □.

Next we prove that ν_θ is an invariant measure. We consider ν_θ as initial state and denote with $S(t)$ the semigroup of $X(t)$. We see that because of Lemma 3.7, we can couple the two initial states $\delta_{\underline{\theta}}$ and $\delta_{\underline{\theta}}S(t)$. Furthermore for a sequence ν_n which has bounded second moments and converges weakly to ν as $n \to \infty$, one has $\nu_n S(t) \Rightarrow \nu S(t)$ as $n \to \infty$. Therefore (recall the classical fact that a successful coupling ensures the weak convergence of the difference measure since by the coupling inequality one gets convergence for uniformly continuous functions)

$$\nu_\theta S(t) = w - \lim_{n\to\infty} \mathcal{L}(X_{t_n+t}) = w - \lim_{n\to\infty} \mathcal{L}(X_{t_n}) = \nu_\theta \tag{3.102}$$

and hence ν_θ is an invariant measure with intensity θ, which is trivially translation invariant.

It remains to show the mixing property of ν_θ. Consider the random field $\{\langle x_\xi, f\rangle\}_{\xi\in\Omega},\ f \geq 0$ under the law ν_θ. For the mixing property of this random field we use the following two facts. (i) The mixing property holds for the total mass process. (ii) The relative weights distribution $\widehat{\nu}_\theta$ is mixing. From there two facts the mixing property is immediate.

The first point was proved in [CG1] based on the representation of the invariant measure as weak limit of the law at time t for $t \to \infty$ and starting in the constant state. The second point is obtained as follows. The coalescent which is for given total mass process dual for the relative weight process (see chapter 1(b)(i) for a definition) has the property that the probability of at least one coalescent event occurring tends to 0 if all distances between initial positions tend to infinity. To see this use the relation (1.86) which implies that the coalescent started in two particles has an expected occupation time tending to zero as the initial positions are moved apart. Having this we have to verify that the positions of the coalescent are such that the corresponding law becomes asymptotically such that the analog of (3.82) holds. However, since the initial state is constant this is now trivial.

Proof of Lemma 3.7 This will be essentially a L_2-calculation. Assume first that under the initial measure $E[\bar{x}_\xi^2] < \infty$. The key point is to show, that for every $f \in C([0,1])$:

(3.103)
$$\sup_{t\geq 0} \left\{ E\left[\langle x_\xi(t), f\rangle \langle x_\eta(t), f\rangle\right] - E\left[\langle x_\xi(t), f\rangle\right] E\left[\langle x_\eta(t), f\rangle\right]\right\} \underset{|\xi-\eta|\to\infty}{\longrightarrow} 0.$$

This is done by explicit calculation for the second moment field, which is based on a calculation with the generator and partial integration for semigroups yielding:

(3.104)
$$\begin{aligned} & E[\langle x_\xi(t), f\rangle\langle x_\eta(t), f\rangle] \\ = & \sum_{\xi',\eta'} a_t(\xi,\xi')a_t(\eta,\eta')E\left[\langle x_{\xi'}(0), f\rangle \langle x_{\eta'}(0), f\rangle\right] \\ & +2\int_0^t \sum_\zeta a_{t-s}(\xi,\zeta)a_{t-s}(\eta,\zeta)E[h_\zeta(\bar{x}_\zeta(s))\langle x_\zeta(s), f^2\rangle]ds \\ = & \sum_{\xi',\eta'} a_t(\xi,\xi')a_t(\eta,\eta')E\left[\langle x_{\xi'}(0), f\rangle \langle x_{\eta'}(0), f\rangle\right] \\ & +2\int_0^t \hat{a}_{2(t-s)}(\xi,\eta)g_f(s)ds, \end{aligned}$$

with

(3.105)
$$g_f(t) = E[h(\bar{x}_\zeta(t))\langle x_\zeta(t), f^2\rangle].$$

If $\nu = \delta_{\underline{\theta}}$, then the r.h.s of (3.104) satisfies the bound

$$\leq \langle\theta, f\rangle^2 + 2\int_0^t \widehat{a}_{t-s}(0,0)g_f(s)ds$$

Focus first on the assertion, about ν_θ. We see with the transience of $\hat{a}$, that $\int_0^\infty \hat{a}_s(0,0)ds < \infty$. Since $g(x) = o(x^2)$ as $x \to \infty$ this implies via a standard renewal argument (see [DG1, 96]) that for $\nu = \delta_{\underline{\theta}}$ (recall (3.100)):

(3.106)
$$sup_{t\geq 0}\left\{Eg_f(x_\xi(t))\right\} < \infty.$$

With the fact that

(3.107)
$$\int_0^\infty \hat{a}_t(\xi,\eta)dt \underset{|\xi-\eta|\to\infty}{\longrightarrow} 0,$$

we obtain then (3.103).

Now turn to the second assertion, the one about finite $t > 0$ i.e. consider a law $\nu S(t)$, with $\nu \in \mathcal{T}_\theta$ and $E_\nu(\bar{x}_\xi^2) < \infty$. We bound the first term in the bound on the r.h.s. of (3.104) using Cauchy-Schwarz by

$$\sum_{\xi', \eta'} a_t(\xi, \xi') a_t(\eta, \eta') E(\bar{x}_\xi)^2 \|f\|_\infty \tag{3.108}$$

which gives together with the argument above that for every $t : \mathcal{L}(X(t)) \in \mathcal{T}_\theta$.

For initial states which are ergodic with intensity θ use the truncated states $(x_\xi(0) \wedge n)_{\xi \in \Omega}$, which are in $\mathcal{T}_\theta$, have bounded second moments, and the fact that they approximate by (3.60) the original process uniformly in t. This implies trivially also the assertion of Lemma 3.7 for $\nu \in \mathcal{T}_\theta$ translation invariant, ergodic with finite intensity measure.

This completes the proof of Theorem 2 part (a), (b).

(ii) Proof of Theorem 2 Part (c) and (d)

(c) Here the simpler part is the transient case, on which we focus first. It remains to show that the functionals $\bar{X}(t)$ and $\widehat{X}(t)$ have laws which converge to $\bar{\nu}_\theta$ and $\widehat{\nu}_\theta$, which are the images of ν_θ under the maps$^-$, respectively, $\widehat{\cdot}$.

This is easy for the total mass process. If we start in the initial state $\delta_{\underline{\theta}}$ the $\bar{x}_\xi(t)$ are uniformly integrable, since the second moments of $\bar{x}_\xi(t)$ remain bounded in time, as is seen from the argument following (3.104). Therefore the weak convergence of $\mathcal{L}(X(t))$ implies the one of $\mathcal{L}(\bar{X}(t))$ as $t \to \infty$ for the special initial state $\delta_{\underline{\theta}}$. By the successful coupling result this generalizes to initial states in $\mathcal{T}_\theta$.

The only issue is the relative weight process. Since we know that $\mathcal{L}(X(t))$ and $\mathcal{L}(\bar{X}(t))$ both converge and since $P(x_\xi(t) = 0) = 0$ for all $t > 0$, the convergence of $\mathcal{L}(\widehat{X}(t))$ follows. Simply use that $x_\xi(t) = \bar{x}_\xi(t)\widehat{x}_\xi(t)$.

(d) The process $\widehat{X}(t)$ is studied by considering its law conditioned on the complete total mass process $(\bar{X}(s))_{s \geq 0}$. For this conditioned law we have identified in Proposition 0.4 the evolution mechanism. Focus now on this conditioned mechanism and use the duality theory of subsection 1(b).

We see immediately from the duality relation of Proposition 1.1 that the quantity

$$\langle \widehat{x}_\xi(t), f^2 \rangle - \langle \widehat{x}_\xi(t), f \rangle^2 = \mathrm{Var}_{\widehat{x}_\xi(t)}(f) \tag{3.109}$$

has an expectation tending to 0 or remaining bounded away from 0 as $t \to \infty$ exactly if $\widehat{G}_t \to \infty$, respectively if $\widehat{G}_t$ is stochastically bounded. Namely the quantity in question determines whether for the coalescent started with two particles the probability of coalescence till time t tends for $t \to \infty$ to 1 or not.

If $h(\bar{x}) \geq c > 0$ then we can bound the coalescence probability away from 0 during each meeting time of the two walks. The result claimed is then immediate.

4. The historical process (Proof Theorem 3 and Theorem 4)

Since the historical process "contains" the ordinary process one might think that in a sense all proofs should be carried out on the level of historical processes. Since on the other hand the dynamics is driven by data available without information on the history, the arguments we gave in sections 2 and 3 are in fact the basic tools needed in the proofs of the theorems on the historical process and the trick is to use those arguments properly on the historical level. We treat in separate subsections 4(a) and 4(b) the existence and uniqueness problem on the one hand and the longtime behavior on the other side. In a third subsection 4(c) we focus on the clan decomposition of the (infinitely old) equilibrium historical process.

(a) Construction of the historical process: Proof of Theorem 3. We begin by sketching our strategy. In the case $h(x) \equiv h$, a constant, our process is a multitype *super random walk* and the result is well-known. Namely the solution of the historical martingale problem (Definition 0.6) is a special case of the historical superprocess (see e.g. Dawson and Perkins ([DP]) or Perkins ([P])). From those techniques the *existence* of X^* is fairly straightforward for general h. Namely the existence of the historical process can be obtained as the limit of rescaled historical branching particle systems. Uniqueness is again much more subtle. However since the argument for uniqueness involves the historical version of the process constructed in Section 2 together with ideas which have appeared elsewhere ([DP]) we only describe the main steps but leave out many lengthy details except for the main new argument Proposition 4.1, crucial for uniqueness.

Here is the plan for subsection 4(a). The uniqueness in the general case is carried out by generalizing the argument in the proof of Theorem 1 in Chapter 2. We first verify that the total mass process in this case can be identified with the total mass process of Chapter 2 (cf. Lemma 2.1) and consider the associated relative weight process which is a time-inhomogeneous interacting Fleming-Viot system. The main new ingredient is to then prove the uniqueness of the one-dimensional marginal distributions of the relative weight process in the historical context and this is done in Proposition 4.1.

The subsection 4(a) has three main parts. We begin by showing the existence of X^* and continue by introducing the pair $(\bar{\mathcal{X}}^*, \widehat{\mathcal{X}}^*)$ (which is the suitable analogue of $(\bar{X}, \widehat{X})$ of chapter 2 and then we show the uniqueness of X^*.

Proof of Existence.
The historical process can be obtained as the diffusion limit of the historical branching particle models $(X^*_a(t))_{t\geq 0}$ and this fact allows us to obtain a càdlág version of our process $(X^*(t))_{t\geq 0}$. For that purpose recall the description of the particle process and its historical process at the very beginning of the paper. The historical process of the particle model can also be characterized by a martingale problem. To do this we proceed exactly as in (0.59) - (0.68) but we have to build in the fact that the migration is still random and hence we have also jumps rather than having for the migration part a deterministic flow given by systems of differential equations. This means that in the increasing process we have an additional term given by the quadratic variation arising due to migration jumps.

A migration jump at time s is given by

$$y \to y'_{s,\xi,\xi'}(u) = \begin{cases} y(u) + (\xi' - \xi) & u \geq s \\ y(u) & u < s. \end{cases} \tag{4.1}$$

This jump occurs at rate $a(\xi', \xi)$, the rate of jumps from ξ to ξ' at time s for the underlying random walk (recall that the particle motion is governed by $\bar{a}$ according to our convention).

Let $(X^*_a(t))_{t\geq 0}$ be a solution of the corresponding well-posed martingale problem. The existence follows immediately since it is a jump process, where the state space $\mathcal{E}^*$ is chosen such that only finitely many jumps occur at each site in finite time intervals.

Now consider for a given state $X(0) \in \mathcal{E}$ the initial states $\lfloor X(0)/\varepsilon \rfloor$. Furthermore give every particle and its marked (with the type in $[0,1]$) path of descent the mass ε and speed up the branching rate per particle by ε^{-1} and scale h to $h(\varepsilon x)$. Denote the resulting process by

$$(X^{*,\varepsilon}_a(t))_{t\geq 0}\,. \tag{4.2}$$

Then we claim that via well-known results we get:

Lemma 4.1.

$$\mathcal{L}\left((X^{*,\varepsilon}_a(t))_{t\geq 0}\right) \underset{\varepsilon\to\infty}{\Longrightarrow} \mathcal{L}\left((X^*(t))_{t\geq 0}\right). \tag{4.3}$$

and $\mathcal{L}\left((X^*(t))_{t\geq 0}\right)$ *is a solution to the martingale problem (0.65) in* $D([0,\infty),\mathcal{M}(D(\mathbb{R},\Omega\times[0,1])))$. □

Proof We first have to prove two facts

- The family $\{\mathcal{L}((X_a^{*,\varepsilon}(t))_{t\geq 0}),\varepsilon>0\}$ is tight as a family of probability laws on $D([0,\infty),\mathcal{M}(D(\mathbb{R},\Omega\times[0,1])))$ as $\varepsilon\to 0$.
- All limit points of the above set as $\varepsilon\to 0$ satisfy the martingale problem (0.68).

Step 1 (h is bounded) If h is bounded above, then the proof of these two assertions is not essentially affected by the fact that our branching rate is not constant and we refer the reader to [DP] Theorem 7.16. The uniqueness proved below establishes convergence and therefore the existence of a unique solution to the martingale problem (0.65) for X^* in $D([0,\infty),\mathcal{M}(D(\mathbb{R},\Omega\times[0,1])))$ in this case.

Step 2 (general case) In the general case the existence of a solution of the martingale problem (0.68) with càdlàg paths, uses a truncation argument to first work with $h_K(x)=h(x)\wedge K$ for which the previous step applies. Then it is an exercise in stochastic analysis to obtain the uniform integrability necessary to verify that the limit points of $P^{(K)}$ as $K\to\infty$ also satisfy the required martingale problem on $D([0,\infty),\mathcal{M}(D(\mathbb{R},\Omega\times[0,1])))$ (See [DG1], [G]). This completes the proof of the Lemma 4.1.

The Pair $(\bar{\mathcal{X}}^*,\widehat{\mathcal{X}}^*)$**.**

We begin by identifying the current total mass and relative weights configurations associated to the historical process, X^*.

Define the *current total masses* at $\xi\in\Omega$:

$$\bar{\mathcal{X}}^*(t)=(\bar{x}^*_\xi(t))_{\xi\in\Omega}, \tag{4.4}$$

with

$$\bar{x}^*_\xi(t)\in\mathcal{M}(D(\mathbb{R},\Omega\times[0,1])),$$

given by

$$\bar{x}^*_\xi(t)=X^*(t)(\{y\in D(\mathbb{R},\Omega\times[0,1])|y(t)\in\{\xi\}\times[0,1]\}). \tag{4.5}$$

Next comes the process of the current relative weights denoted

$$\widehat{\mathcal{X}}^*(t)=(\widehat{x}^*_\xi(t))_{\xi\in\Omega} \tag{4.6}$$

and defined as follows. For every $A\in\mathcal{B}(D(\mathbb{R},\Omega\times[0,1]))$ put:

$$\begin{aligned}\widehat{x}^*_\xi(t)(A) \;=\; & [X^*(t)(A\cap\{y\in D(\mathbb{R}),\Omega\times[0,1])|y(t)\in\{\xi\}\times[0,1]\})] \\ & [\bar{X}^*(t)(\{y\in D(\mathbb{R}),\Omega\times[0,1])|(\pi_1 y)(t)=\xi\})]^{-1}.\end{aligned} \tag{4.7}$$

We have now defined the bivariate process $(\bar{\mathcal{X}}^*(t),\widehat{\mathcal{X}}^*(t))_{t\geq 0}$ replacing $(\bar{X}(t),\widehat{X}(t))_{t\geq 0}$ from section 2.

Proof of Uniqueness

As before in chapter 2 we prove first the uniqueness of $\bar{\mathcal{X}}^*$ and then the uniqueness of the process $\widehat{\mathcal{X}}^*$ for given $\bar{\mathcal{X}}^*$. In both steps we make heavy use of chapter 2.

The first observation is that the total mass process is uniquely determined by showing:

Lemma 4.2. *The process* $\bar{\mathcal{X}}^*(t)$ *solves the* $(\bar{G},\delta_{\bar{x}})$*-martingale problem.* □

Proof We simply insert in the martingale problem for $X^*(t)$, a special subclass of functions for which G^* is defined namely functions of the type $\Phi(t,y)$, which depends up to time t_0 only on the location component of the endpoint of the path, namely $y(t)$, i.e. is of the form (π_1 is projection on the space component Ω of the path)

$$\Phi(t,y)=g_0((\pi_1 y)(t\wedge t_0)),\quad g_0\in C(\Omega),\quad t_0\in\mathbb{R}, \tag{4.8}$$

to recover the $(\bar{G}, \delta_{\bar{x}})$-martingale problem on $[0, t_0]$ by explicit calculation.

Next we prove the uniqueness of the solutions to the martingale problem for the relative weights process $\hat{\mathcal{X}}^*$ given $\bar{\mathcal{X}}^*$ and also the solution of the full historical martingale problem (0.65). The proof uses the uniqueness for the finitely many types martingale problem proved in Chapter 2. Recall the terminology in Definition 0.5.

Proposition 4.1. *(a) For given total mass process $(\bar{\mathcal{X}}^*(t))_{t\geq 0}$ the law of the process $(\widehat{\mathcal{X}}^*(t))_{t\geq 0}$ is up to equivalence uniquely determined and the law of the latter is a measurable function of $\bar{\mathcal{X}}^*$.*
(b) The law of the process defined by

$$X^*(t)(A \cap \{y|y(t) \in \{\xi\} \times [0,1]\}) = \bar{x}^*_\xi(t)[\widehat{x}^*_\xi(t)(A)], \qquad \text{for all } A \in \mathcal{B}(D(\mathbb{R}, \Omega \times [0,1])), \xi \in \Omega, t \geq 0, \tag{4.9}$$

gives the unique solution of the martingale problem (0.68) in $C([0,\infty), \mathcal{E}^*)$. □

Proof of Proposition 4.1
The proof proceeds in five steps.

Step 1 Our goal is to construct the finite dimensional distributions of $(\widehat{\mathcal{X}}^*(t))_{t\geq s}$ for a given starting point $\widehat{\mathcal{X}}^*(s)$ and show they are unique. In other words we determine and construct for $s \leq t_1 < t_2 < \cdots < t_m, m \in \mathbb{N}$:

$$\mathcal{L}[\widehat{\mathcal{X}}^*(t_1), \ldots, \widehat{\mathcal{X}}(t_m)|(\bar{\mathcal{X}}^*(t))_{t\geq 0}]. \tag{4.10}$$

We begin with the case $m = 1$. In other words we are going to construct and determine now in the next step for a fixed value $t > 0$:

$$\mathcal{L}[\widehat{\mathcal{X}}^*(t)]. \tag{4.11}$$

That is, we obtain for every initial state and initial time a unique law.

Next note that by (4.11) we obtain already a unique process, since the coefficients in the martingale problem depend only on the current state, so that it suffices to show that starting in every point of the state space the one-dimensional marginals are uniquely determined for all times. Hence it suffices to consider $m = 1$.

Step 2 The first idea is to determine the law of $\mathcal{X}^*(t)$ by considering the projection of this random measure on the *path space* $D(\mathbb{R}, \Omega \times [0,1])$ onto a random measure depending only on the observation of the path at a *finite number of time points.*

The second idea is to use as the key ingredient a generalization of the construction in chapter 2, namely we work with a modified type space, which encodes the type in the sense of the process $X^*(t)$ and in addition the locations of a tagged mass at finitely many fixed earlier time points. This means that the type space $[0,1]$ will be replaced by $\Omega^k \times [0,1]$. We fix the starting time s of the historical process at this point.

Denote by $\tau = \tau(n)$ a collection

$$\tau = \{s < t_1 < \cdots < t_n < t_{n+1} = t\} \tag{4.12}$$

and define ($\sigma(E)$ is the σ-algebra generated by E):

$$\mathcal{B}_\tau = \sigma(\{(y(t_1), \ldots, y(t_{n+1}))|y \in D(\mathbb{R}, \Omega \times [0,1])\}). \tag{4.13}$$

We define for fixed t:

$$\widehat{\mathcal{X}}^{*,\tau}(t) = \widehat{\mathcal{X}}^*(t)|_{\mathcal{B}_\tau}. \tag{4.14}$$

This means we have to show that for every partition τ of $[0,t]$:

$$\mathcal{L}[\widehat{\mathcal{X}}^{*,\tau(n)}(t)] \quad \text{is uniquely determined and a measurable function of } \bar{\mathcal{X}}^*. \tag{4.15}$$

To prove (4.15) we apply the results of chapter 2 to state spaces of the form $\Omega^k \times [0,1]$ with $k \in \mathbb{N}$ instead of $[0,1]$. This is no problem since for every k we can embed $(\Omega^k \times [0,1])$ into $[0,1]^2$ and then we only have to observe that what we used about $[0,1]$ was the existence of a countable dense subset in $C([0,1])$. Obviously we have that as well for $C([0,1]^2)$.

Step 3 In this step we shall prove (4.15) by identifying inductively the law in terms of laws of certain time-inhomogeneous Fleming-Viot processes. Inductive means here that for a given $\tau = \tau(n)$ we consider successively

$$\tau^1 = \{t_1, t_2\} \quad \tau^2 = \{t_1, t_2, t_3\}, \; \dots \; , \tau^{(n)} = \{t_1, \dots, t_{n+1}\}, \quad t_1 = s, \quad t_{n+1} = t \tag{4.16}$$

and define this way the law of $\widehat{\mathcal{X}}^{*,\tau(n)}(t)$. This arises since we incorporate in the type the information on the position of paths at a given set of earlier time points.

Define the processes (here $Y^{k,k+1} = (y_\xi^{k,k+1})_{\xi \in \Omega}$, with $y_\xi^{k,k+1} \in \mathcal{P}(\Omega^k \times [0,1])$):

$$(Y^{(k,k+1)}(t))_{t \geq t_k}, \tag{4.17}$$

to be processes starting in initial states determined below in (4.19) which are solutions of the conditional martingale problem (0.51) with the rates specified in (0.57) and (0.58) but with type space:

$$\Omega^k \times [0,1]. \tag{4.18}$$

We showed in chapter 2 that there exists a unique solution in the equivalence class defined in (2.7).

From these objects we shall define $\mathcal{X}^{*,\tau(n)}$ by the following recursive scheme which gives the initial states of the above processes. Define a sequence $(\widehat{Z}^k)_{k \in \mathbb{N}}$ of $\mathcal{P}(\Omega^k \times [0,1])$-valued random measures together with the processes $(Y^{k,k+1}(t))_{t \geq 0}$ by setting: (here ξ plays the role of the new type ω_{k+2} added in this step)

$$\begin{aligned} \widehat{Z}_\tau^{k+1}(t_{k+1})((\omega_1, \dots, \omega_{k+1}, \xi; A)) &= y_\xi^{(k,k+1)}(t_{k+1})((\omega_1, \dots, \omega_{k+1}; A)), \\ & \qquad \xi \in \Omega, \quad A \in \mathcal{B}([0,1]) \\ Y^{k,k+1}(t_k) &= \widehat{Z}_\tau^k(t_k), \\ \widehat{Z}_\tau^0((\xi; A)) &= \widehat{x}_\xi(s)(A), \quad \xi \in \Omega, \quad A \in \mathcal{B}([0,1]). \end{aligned} \tag{4.19}$$

Step 4. In this step we now prove that (recall $t_{n+1} = t$):

Lemma 4.3. *For every realization of* $(\bar{\mathcal{X}}(t))_{t \geq 0}$*:*

$$\mathcal{L}\left(\widehat{Z}_\tau^{n+1}(t)\right) = \mathcal{L}\left(\widehat{\mathcal{X}}^{*,\tau(n)}(t)\right), \tag{4.20}$$

and consequently for the unconditioned quantities:

$$\mathcal{L}\left(\bar{\mathcal{X}}^* \cdot \widehat{\mathcal{X}}^{*,\tau(n)}(t)\right) = \mathcal{L}\left(X^*(t)|_{\mathcal{B}_{\tau(n)}}\right). \qquad \square \tag{4.21}$$

We introduce the following notation, namely we associate with the $\mathcal{L}(\widehat{Z}^{n+1})$ the probability law, P^τ, of a random measure on $D(\mathbb{R}, \Omega \times [0,1])$ by taking the piecewise constant functions with the values at times $\{t_1, \dots, t_{n+1}\}$ defined by the sequence defined in (4.19).

At this point we are going to make heavy use of our previous results in chapter 2. We return to the martingale problem of the historical process, which was given in (0.68). We consider in the algebra $\tilde{\mathcal{A}}$ of functions used there, a function of the type

$$\Phi(t, y) = \prod_{j=1}^{n+1} g_j(t, y(t \wedge t_j)) \tag{4.22}$$

with $t_1, t_2, \dots, t_n$ given by τ and with $t_{n+1} = t$. We write $y_\tau = (y(t_1), \dots, y(t_{n+1}))$ and $\Phi(t, y_\tau)$ since $\Phi(t, y)$ depends only on the values in $t_1, \dots, t_{n+1}$, at time $t = t_{n+1}$.

Then we have to show that the expectation of $\langle \Phi, \widehat{\mathcal{X}}^*(t) \rangle$ conditioned on $(\bar{\mathcal{X}}^*(s))_{s \geq 0}$ can be calculated by the random variables $Z_\tau^{|\tau|}$, i.e. by $\int \Phi(t, y_\tau) Z_\tau^{|\tau|}(t)(dy_\tau)$, this means precisely:

$$E[\langle \Phi(t, \bullet), \mathcal{X}^{*,\tau}(t)(\bullet) \rangle] = E[\langle \Phi(t, \bullet), Z_\tau^{|\tau|}(t)(\bullet) \rangle], \quad \text{for all } \Phi \in \mathcal{A}^*. \tag{4.23}$$

This is now proved by induction over n. We assume that the assertion holds for $\mathcal{B}_{\tau(n)}$ for all $\Phi(t,\bullet)$ from our collection of test functions and we conclude that this then holds for $\mathcal{B}_{\tau(n+1)}$. We first condition on $\mathcal{B}_{\tau(n)}$, since then

$$E\left[\langle\Phi(t,\bullet),\widehat{\mathcal{X}}^*(t)\rangle|\mathcal{B}_{\tau(n)},(\bar{\mathcal{X}}^*(t))_{s\geq 0}\right] = \prod_{j=1}^{n} g_j(t,y(t_j))\cdot E\left[\langle g_{n+1}(t,y(t)),\widehat{\mathcal{X}}^*(t)(dy)\rangle|\mathcal{B}_{\tau(n)},(\bar{\mathcal{X}}^*(s))_{s\geq 0}\right]. \tag{4.24}$$

We have to show that the expectation on the r.h.s. above equals (here $\lambda\in\Omega^{n+2}\times[0,1]$) for fixed $(\bar{\mathcal{X}}^*(s))_{s\geq 0}$:

$$E\left[\langle g_{n+1}(t,\bullet),\ \widehat{Z}_\tau^{n+1}(d\bullet)\rangle|\widehat{Z}_\tau^n\right]=:\tilde{g}_n(t,\widehat{Z}_\tau^n). \tag{4.25}$$

Once we have this relation we can proceed by iterating this procedure to obtain the claim (4.20). Namely we observe that the distribution of $\widehat{\mathcal{X}}^*(t)|_{\mathcal{B}_{\tau(n)}}$ for given realization of $(\bar{\mathcal{X}}^*(t))_{t\geq 0}$ equals the one for $(\widehat{Z}_\tau^1,\dots,\widehat{Z}_\tau^n)$. This gives the claim for $n+1$.

The relation (4.25) now comes from the martingale problem for the historical process applied in the time interval $[t_n,t_{n+1}]$ by stochastic calculus. We observe first from the form of the generator of the martingale problem (that is a branching Markov path process) that the object $\mathcal{X}^*(t)$ induces a measure ϱ on the space of path which has the following property. Denote by $\mathcal{B}_I$ the sub-σ-algebra (of the Borel algebra of $D(\mathbb{R},\Omega)$) generated by the path in the time interval I. Then

$$\varrho(A),\varrho(B) \text{ are conditioned on } \mathcal{B}_{t_n} \text{ conditionally independent, where } A\in\mathcal{B}_{[0,t_n]}, B\in\mathcal{B}_{(t_n,t]}. \tag{4.26}$$

This implies that during one time step the expectation on the r.h.s. of (4.24) can be expressed in terms of the multi-type state dependent branching process with type space $\Omega^n\times[0,1]$ itself. Knowing this we have to see only that the dynamics of the relative weights is as claimed that is given by the process in (4.17). Then however we are back to chapter 2. If we put first $g(s,y)=\tilde{g}(y)$for all $s\in[0,t]$, then we see immediately from the expression (0.64) that we have the exact same structure as we had in chapter 2 only the type space has changed. The inclusion of the t-dependent g_{n+1} is straightforward.

Step 5 In this step we verify the convergence: if we choose any sequence $(\tau(n))$ of partitions $(t_1^{(n)},\dots,t_{n+1}^{(n)})$ of $[0,t]$ which are successive refinements of each other and with $\sup_{i=0,\dots,n}|t_{i+1}^{(n)}-t_i^{(n)}|\to 0$ as $n\to\infty$, then the sequence

$$P^{\tau(n)} \tag{4.27}$$

as laws on $D(\mathbb{R},\Omega\times[0,1])$ converges to a limit, which is independent of the choice of $(\tau(n))_{n\in\mathbb{N}}$, that is,

$$P^{\tau(n)}\underset{n\to\infty}{\Longrightarrow}\mathcal{L}[X^*(t)]. \tag{4.28}$$

Here the key point is the fact that by the prescription given in the previous step and the well-posedness of the martingale problem for the Y-processes used there, we obtain a consistent family of laws, $\{\mathcal{L}(\widehat{\mathcal{X}}^{*,\tau(n)}(t)),n\in\mathbb{N}\}$, on

$$(\mathbb{R})^{\Omega\times[0,1]}. \tag{4.29}$$

To prove that the limiting random measure can be realized on $D(\mathbb{R},\Omega\times[0,1])\subseteq(\mathbb{R})^{\Omega\times[0,1]}$ it remains to show that the family of laws, which is given by:

$$\left\{P^{\tau(n)}\right\}_{n\in\mathbb{N}} \tag{4.30}$$

is tight in $\mathcal{P}(D(\mathbb{R},\Omega\times[0,1]))$. In [D], Lemma 3.2.8 it is shown that we have tightness if the mean measures are tight. The mean measures however are given by the laws of the $\{t_1,\dots,t_{n+1}\}$ piecewise constant approximations to the $(a_s^t)_{s\leq t}$ random walk. But it is easy to verify that the later is tight in $\mathcal{P}(D(\mathbb{R},\Omega\times[0,1]))$. The above relations together prove that $X^*(t)$ is uniquely determined and can be realized as a random measure on $D(\mathbb{R}, \Omega \times [0,1])$. This completes the proof of Proposition 4.1.

(b) The longtime behavior of the historical process: Proof of Theorem 4. In this section we prove Theorem 4 by extending the coupling argument used in Section 3 to prove Theorem 2. Roughly we shall consider the *historical process associated with our coupled dynamics* we constructed in section 3(a). The lifting of the coupling argument of part (a) onto a historical process level is however a bit trickier than lifting the construction of the process to the one of the historical process. The reason is that individuals in the bivariate process are now marked by path and type and the combination (type, paths) can change in time and furthermore the set of paths on which the historical process is concentrated is not countable. However we shall adapt the coupling argument to both problems and we shall be able to derive, by making directly use of the coupling argument of part (a), that also our historical coupling is successful. The key step, Lemma 4.4 below, is based on the main feature our coupled dynamics of chapter 3, namely that coupled particles follow the same path. Based on this fact we utilize the successful coupling property at the particle level proved in Section 3. We proceed in steps.

Step 1 In this preliminary step we introduce the appropriate notion of the coupled dynamics and a distance between probability measures on the state space of this coupled process. Denote with $X_t^{*,i}$ $i = 1, 2$ the two historical processes we want to compare and which started at time 0 in two given random configurations.

Consider a process $Z^*(t)$ with values in the measures on the space of pairs of coupled marked path, which is given by

$$\mathcal{M}((D(\mathbb{R}, \Omega) \times \{[0,1] \cup \{*\}\})^2) = \mathcal{M}(\widetilde{E}^*) \tag{4.31}$$

and denote with π_1, π_2 the projections from $\widetilde{E}^*$ to its first respectively second component and $\bar{\pi}_1, \bar{\pi}_2$ the corresponding maps acting on measure on $\widetilde{E}^*$. We need to choose the evolution of $(Z^*(t))_{t \geq 0}$ such that

$$\begin{aligned} \mathcal{L}\big((\bar{\pi}_1 Z_t^*)_{t\geq 0}\big) &= \mathcal{L}\left((X_t^{*,1})_{t\geq 0}\right) \\ \mathcal{L}\big((\bar{\pi}_2 Z_t^*)_{t\geq 0}\big) &= \mathcal{L}\left((X_t^{*,2})_{t\geq 0}\right). \end{aligned} \tag{4.32}$$

In addition the process Z_t^* should be "close to the diagonal" for $t \to \infty$. We shall next define "closeness to the diagonal" in the new path-valued context.

As before it is more convenient to *start the processes at time $-t$ and observe them at time 0.* We use this convention in the next two definitions.

We begin by looking at paths with values in Ω. We will base our notion of distance of two states of X^*, i.e. elements of $\mathcal{M}(D(-\infty, \infty), \Omega)$ on the same principles as in part (a) step 1. Since relevant elements in $\mathcal{M}(D(\mathbb{R}, \Omega))$ are only locally finite we focus first on the measure restricted to the set of all paths which at time 0 (recall we start in $-t$) are in the site ξ. We call this set $D_\xi(\mathbb{R}, \Omega)$. Note that we cannot expect to couple arbitrarily long time stretches of the paths but only *finite end pieces of the ancestral paths.* Hence fix a number $T \in [0, \infty)$ and define for every T the *distance of two paths* $y^1, y^2 \in D(\mathbb{R}, \Omega)$ by (note this is strictly speaking only a quasi distance, which might be 0 for different paths):

$$d_{T,\xi}(y^1, y^2) = 1 - 1\!\!1(y^1(s) = y^2(s) \quad \forall s \in [-T, 0]), \quad \text{for } y^1, y^2 \in D_\xi(\mathbb{R}, \Omega). \tag{4.33}$$

We lift this distance function to a *distance function* $\widetilde{d}_T$ between two measures on $D(\mathbb{R}, \Omega)$:

$$\widetilde{d}_{T,\xi} : \mathcal{M}((D(\mathbb{R}, \Omega))^2) \to \mathbb{R}^+, \tag{4.34}$$

which is obtained from(3.6) if we replace in that formula the interval $[0,1]$ by the path space $D_\xi(\mathbb{R}, \Omega)$ and $d(\cdot, \cdot)$ by $d_{T,\xi}(\cdot, \cdot)$.

We incorporate next the type into the distance function. We set for path-type combinations $(y_1, u), (y_2, v)$ with two paths y_1, y_2 which occupy at time 0 the point ξ (recall (4.33) for $d_{T,\xi}$ and (3.6) for d):

$$d^*_{T,\xi}\left((y_1, u), (y_2, v)\right) = d_{T,\xi}(y_1, y_2) + d(u, v) \quad \text{for } y_i \in D_\xi(\mathbb{R}, \Omega);\ u, v \in [0,1] \cup \{*\}. \tag{4.35}$$

As before this induces a Vasserstein quasi-metric $\widetilde{d}^*_{T,\xi}$ for elements of $\mathcal{M}(\widetilde{E}^*)$.

Now the key quantity for the comparison of the two processes $\mathcal{X}^{*,i}, i = 1, 2$:

$$\Delta^*_{T,\xi}(t) = E\left[\tilde{d}^*_{T,\xi}(\bar{\pi}_1 Z^*_t,\ \bar{\pi}_2 Z^*_t)\right]. \tag{4.36}$$

We shall study this quantity again for the case where θ has finite support. Our goal will be to show that for suitable initial states $\Delta^*_{T,\xi}(t) \to 0$ as $t \to \infty$ for all $T \geq 0$, that is we have a successful coupling. Precisely:

Proposition 4.2. *Assume that θ is an atomic measure with finitely many atoms and that $\mathcal{L}(X^{*,1}(-t))$ and $\mathcal{L}(X^{*,2}(-t))$ are concentrated on constant paths and the collection of the projections on the components of path being at ξ at time 0 with ξ running through Ω are both translation invariant and ergodic with $E(X^{*,I}(\{y|y(0) \in \{\xi\} \times [0,1]\})) = \theta, \quad i = 1, 2$. Then*

$$\Delta^*_{T,\xi}(t) \underset{t \to -\infty}{\longrightarrow} 0, \quad \forall \xi \in \Omega,\ T \geq 0. \quad \square \tag{4.37}$$

This result then completes the proof of Theorem 4 part (b) for θ atomic. The case of non-atomic θ follows again as in section 3 (3.95) - (3.99) we don't repeat these straightforward details.

This leaves us now with the following two tasks to complete the proof. We have to define the dynamics of $(Z^*(t))_{t \geq 0}$, which induces the coupled dynamics and then to verify (4.37). We shall do this by first defining $(Z^*_a(t))_{t \geq 0}$ for a particle model in step 2 for which we then show the assertion analogue to (4.37) in step 3. After that we define in step 4 $\mathcal{L}[(Z^*(t))_{t \geq 0}]$ as a limit point after taking the diffusion limit of the particle model and we shall point our why the argument that the coupling is successful, i.e. $\Delta^*_{T,\xi}(t)$ tends to 0 as $t \to \infty$, carries over from the particle model Z^*_a to Z^*. Here are now the details.

Step 2 Recall at this point the construction of a coupling in section 3(a). There we constructed the coupled dynamic $(Z_a(t))_{t \geq 0}$ by specifying the generator by means of a table giving the possible transitions and the associated rates. These rates however were derived by thinking of a process with individual particles for which we specified some simpler transition rules, which then translated into the generator once we do not observe individual particles but numbers of particles with a certain type and a certain location. Instead of writing down again the tables we now only give the new rules on the level of individuals and their ancestral path and explain how to derive from that the new transition table induced on the random measures on ancestral path, only for one of the transitions.

We start now with individual particles which migrate according to $\bar{a}(\cdot,\cdot)$ and branch according to the branching rate h. Each individual has a mark of the form

$$(y, u) \quad y \in D(\mathbb{R}, \Omega), \quad u \in [0,1], \tag{4.38}$$

where u is the type and y the path of descent of this individual, which is the path from the ancestor at time 0 through its descendents ending at the present (time t) position of the individual. Recall we always continue paths of descent at time t beyond t and before 0 by using the continuation as a constant.

For the dynamic $Z^*_a(t)$ we consider pairs of matched individuals characterized by pairs of marks. We have to deal again with different total numbers of particles at a given site. Therefore we have for the coupled dynamic pairs of matched individuals with marks allowing for an extra type *, this means a single pair of two particles carries a marker:

$$((y_1, u), (y_2, v)) \quad y_1, y_2 \in D(\mathbb{R}, \Omega);\ u, v \in [0,1] \cup \{*\}. \tag{4.39}$$

Let us specify the initial state and the dynamic.

- As *initial state* we choose the state which has the property that all path of descent are constant and equal to the location at time $t = 0$. The matching of the two populations in the beginning is exactly the one given in step 4 in chapter 3.

The dynamics of the process is given by the following rules:

- As long as the two individuals are matched, the paths y_1, y_2 of these two matched individuals follow the *same* path process which is the one induced as explained in subsection 0(d) by the continuous time $\bar{a}(\cdot,\cdot)$-random walk. Whenever a migration transition has been carried out a rematching of individuals takes place instantaneously in order to minimize the distance between the u, v part of the mark of an individual. The rematching is exactly the one given in section 3 and the paths y_2, y_2 are ignored for that matter.
- A birth or death occurs in both parts of $((y_1,u),(y_2,v))$ at the same time at rate $h(\bar{z}^1_\xi(t)) \wedge h(\bar{z}^2_\xi(t))$, where $\bar{z}^1_\xi, \bar{z}^2_\xi$ denotes the number of individuals with type in $[0,1]$ in the first respective second component of the ξ-component of $Z^*_a(t)$ projected from paths to positions at time t. The path till the birth time and the type of a newborn particle is always inherited from the parent particle.
- A birth of a pair
$$((y_1,*),(y_2,v)) \text{ respectively } ((y_1,u),(y_2,*))$$
occurs in addition at rates $(h(\bar{z}^1_\xi(t) - h(\bar{z}^2_\xi(t)))^+, (h(\bar{z}^1_\xi(t) - h(\bar{z}^2_\xi(t)))^-$. Instantaneously a rematching, if possible, takes place, in order to lower the distance between the second components of the type. This rematching is exactly as in subsection 3(a), the paths y_1, y_2 are simply ignored for this transition.

If we now only observe the number of pairs of a certain type $((y_1,u)(y_2,v))$ at a site ξ at time t we obtain a measure valued state and its evolution is a Markov process with values in $\mathcal{M}((D(\mathbb{R},\Omega)\times([0,1]\cup\{*\}))^2$ which shall be denoted:

$$(Z^*_a(t))_{t\geq 0}. \tag{4.40}$$

Note that this process is the historical process associated with the process $(Z_a(t))_{t\geq 0}$. This process can be characterized by its generator, which collects the previous collection of transitions and rates of the individual particles. We refer the reader to section 3 for this process of translation.

As an example note that in $(Z^*_a(t))_{t\geq 0}$ the transition

$$Z \to Z + \delta_{((y_1,u),(y_2,u))} \quad u \in [0,1];\ y_1, y_2 \in D(\mathbb{R},\Omega) \quad \text{with } y_1(t) = y_2(t) = \xi \tag{4.41}$$

occurs at time t at the rates

$$[h(\bar{z}^1_\xi(t)) \wedge h(\bar{z}^2_\xi(t))] \cdot [Z^*_a(t)(\{(y_1,u),(y_2,v))|y_1(t) = y_2(t) = \xi, v = u\})], \tag{4.42}$$

with $\bar{z}^1_\xi = Z_a(\{((y_1,u),(y_2,v))|y_1(t) = y_2(t) = \xi,\ u \in [0,1], v \in [0,1]\cup\{*\}\})$ and $\bar{z}^2_\xi$ is defined by exchanging the role of u and v.

Step 3 For the dynamics $(Z^*_a(t))_{t\geq 0}$ we define the mean distance from the diagonal as induced by the function $\tilde{d}^*_{T,\xi}(\cdot,\cdot)$ from (4.34) as ($\bar{\pi}_1, \bar{\pi}_2$ as the projections on the first respectively second component of Z^*_a):

$$\Delta^{*,a}_{T,\xi}(t) = E[\tilde{d}^*_{T,\xi}(\bar{\pi}_1 Z^*_a(t),\ \bar{\pi}_2 Z^*_a(t))], \tag{4.43}$$

as we defined $\Delta^*_{T,\xi}(t)$ for a diffusion process on the same state space. We now prove

Lemma 4.4. *If $\mathcal{L}[X^i(0)], i = 1,2$ are translation invariant shift ergodic with the same atomic intensity measure θ, then:*

$$\Delta^{*,a}_{T,\xi}(t) \underset{t\to\infty}{\longrightarrow} 0. \qquad \square \tag{4.44}$$

Proof In order to bring the results of chapter 3 into play we observe that by construction of Z^*_a (this can be checked by explicit calculation with the generator):

$$\mathcal{L}\left[(Z^*_a(t)(\{(y_1,u),(y_2,v)),\ y_1(t) = y_2(t) = \xi\})\,;\quad (u,v) \in [0,1]\cup\{*\}\right] = \mathcal{L}[z_\xi(t)]; \quad \forall \xi \in \Omega. \tag{4.45}$$

where $Z_a(t) = (z_\xi(t))_{\xi\in\Omega}$ is the dynamics defined in step 4 in chapter 3. This implies that:

$$\begin{aligned} R_t(\xi) &= E[Z^*_a(t)(\{((y_1,u),(y_2,v))|u \neq v, \quad y_1(t) = y_2(t) = \xi\})] \\ &= E[z_\xi(t)(\{(u,v)|u \neq v\})]. \end{aligned} \tag{4.46}$$

Since we proved in subsection 3 that as $t \to \infty$ the states of $(Z_a(t))_{t\geq 0}$ concentrate on the diagonal as $t \to \infty$ we know that:

$$R_t(\xi) \underset{t\to\infty}{\longrightarrow} 0. \tag{4.47}$$

Finally consider the set (here $J = (D(\mathbb{R}, \Omega) \times ([0,1] \cup \{*\}))^2$):

$$D_1 = \{((y_1, u), (y_2, v)) \in J | u = v \in [0,1]\}. \tag{4.48}$$

We define for the process $(Z_a^*(t))_{t\geq 0}$ the functional $N_{t,\xi}(D_1)$:

$$N_{t,\xi}(D_1) = \sharp\{s \in [0,t] | Z_a^*(s)(D_1) - Z_a^*(s_-)(D_1) \leq -1,\ z_\eta(s) = z_\eta(s_-) \quad \forall \eta \neq \xi\} \tag{4.49}$$

i.e. these quantities measure the jumps where in the process $Z_a^*(t))_{t\geq 0}$ a jump of a unit mass away from D_1 in colony ξ occurs before time t, which is not due to a migration out of ξ. We know by construction that

$$E[N_{t,\xi}(D_1) - N_{s,\xi}(D_1)] = \int_s^t |h(\bar{z}_\xi^1(u)) - h(\bar{z}_\xi^2(u))| ds. \tag{4.50}$$

As a consequence of the result (3.72) in chapter 3 we know for all $T > 0$ that:

$$E[N_{t,\xi}(D_1) - N_{t-T,\xi}(D_1)] \underset{t\to\infty}{\longrightarrow} 0. \tag{4.51}$$

This means that for large t all particles with $u = v \in [0,1]$ in ξ at time $t - T$ keep this property furthermore up to time t. In formulas this reads as follows. Define:

$$\begin{aligned} \widetilde{\Delta}^{*,a}_{T,\xi}(t) \quad = \quad & E[Z_a^*(t)(\{(y_1,u),(y_2,v))|y_1(t) = y_2(t) = \xi, y_1(u) \neq y_2(u) \quad \text{for some } u \in [t-T,t], \\ & \text{or } u \neq v \text{ or } u = *, \text{ or } v = *\})] \underset{t\to\infty}{\longrightarrow} 0. \end{aligned} \tag{4.52}$$

On the other hand by definition of $d(\cdot,\cdot)$:

$$\Delta^{*,a}_{T,\xi}(t) \leq 3\widetilde{\Delta}^{*,a}_{T,\xi}(t). \tag{4.53}$$

This proves (4.44) with (4.52).

Step 4 Now we want to pass for our coupled historical particle process $(Z_a^*(t))_{t\geq 0}$ to the diffusion limit. For this purpose we give every individual the mass ε, blow up the initial configurations by ε^{-1} and we speed up the rates per particle which are based on h by setting (x denotes the number of particles in the colony):

$$h^\varepsilon(x) = \varepsilon^{-1} h(x\varepsilon). \tag{4.54}$$

The resulting processes we call

$$(Z^*_{a,\varepsilon}(t))_{t\geq 0}. \tag{4.55}$$

Clearly for each $\varepsilon > 0$ the process has the property given in Lemma 4.4. If we can find a limit dynamics $(Z^*(t))_{t\geq 0}$ preserving the latter property we are done. Therefore the point now is to show two things:

$$\mathcal{L}[(Z^*_{a,\varepsilon}(t))_{t\geq 0}] \quad \text{is tight for } \ \varepsilon \to 0, \tag{4.56}$$

$$\Delta^{*,a,\varepsilon}_{T,\xi}(t) \underset{\varepsilon\to\infty}{\longrightarrow} \Delta^*_{T,\xi}(t) \tag{4.57}$$

uniformly in $t \geq 0$ along every sequence giving convergence in (4.56). Both these relations complete immediately the proof of Proposition 4.2.

Start with relation (4.56). From subsection 4(a) we know that the two component processes

$$(Z^*_{a,\varepsilon}(t)((\cdot,\cdot),\ D(\mathbb{R},\Omega) \times [0,1]))_{t\geq 0}, \quad (Z^*_{a,\varepsilon}(t)(D(\mathbb{R},\Omega) \times [0,1], (\cdot,\cdot)))_{t\geq 0} \tag{4.58}$$

define a sequence of tight laws. This is an immediate consequence of Lemma 4.1. What we need is tightness of the full process.

However we know that integrating out over the paths gives a tight set according to chapter 3 and furthermore the paths do *not* enter into the matching mechanism. Now it is elementary to check the tightness with the standard criterion using the local characteristics. (Compare the construction in [DGV] section 1. Use test functions on $(D(\mathbb{R},\Omega)\times([0,1]\cup\{*\}))^2$ which have the form $F_1\cdot F_2$ where F_1 depends on the $[0,1]\cup\{*\}$ part of the marker and F_2 on the $D(\mathbb{R},\Omega)$-part of the marker.)

In order to prove (4.57) proceed as follows. The quantity $\Delta^{*,a,\varepsilon}_{\xi,T}$ can be estimated, as we saw in (4.53) in terms of $\widetilde{\Delta}^{*,a,\varepsilon}_{T,\xi}$. These are quantities which depend only on $(Z_a(t))_{t\geq 0}$ via two functionals which we already studied in chapter 3, namely $\Delta^a_\xi(t), E\int_{t-T}^t |h(\bar z^1_\xi(s))-h(\bar z^2_\xi(s))|ds$. It is easy to see that the differential equations ruling those quantities are independent of ε and therefore we get the desired uniform convergence since the total mass of paths ending in ξ is a uniformly (in ε) integrable random variable.

(c) The family decomposition of the equilibrium historical process. In this section we prove Proposition 0.5.

Part (a) Recall that a family is defined to be the mass corresponding to a set of paths having a common initial segment. We know from Lemma 0.1 that during any positive time the dynamics of the historical process leads to a state with countable support. Therefore for reasons of consistency the number of families must be countable. Therefore we can view the equilibrium historical family decomposition as the decomposition of the equilibrium mass on $D(\mathbb{R},\Omega)\times[0,1]$ into subpopulations having a common path in the past from $-\infty$ to some finite time, that is an equivalence class under $\sim$.

Part (b) We must now show that such an equivalence class has only one type and thus corresponds to a unique equivalence class under $\approx$. Consider a family which is multitype. The first observation is that conditioned on the total mass process the evolution of the relative proportions of the different types, is a time-inhomogeneous Fleming-Viot system. Hence it suffices to establish that this Fleming-Viot process degenerates to a single type in finite time.

Now consider a (non-null) family with common historical path $\xi(s):-\infty<s\leq t$. Let $(y_\xi(r))_{r\geq 0}$ be the time-inhomogeneous Fleming-Viot process with resampling rate $h(\bar x_{\xi(r)}(r))/\bar x_{\xi(r)}(r)$ at time $r\in[t_0,t]$ and initial distribution $\mu_{t_0}\in\mathcal{M}_1[0,1]$ at time t_0.

If

$$\lim_{t_0\to-\infty}\int_{t_0}^t h(\bar x_{\xi(s)}(s))/\bar x_{\xi(s)}(s)ds=\infty,\text{except for a } \tilde X^*-\text{null set} \tag{4.59}$$

then it follows (cf. Dawson and Hochberg (1982), Lemma 6.5) that

$$\lim_{t_0\to-\infty} P\left[\left(y_{\xi(t)}(t)=\delta_u \text{ for some } u\in[0,1]\right)\right]=1. \tag{4.60}$$

It remains to verify that (4.59) holds except for an $\tilde X^*$-null set. To do this note that $E(\tilde X^*(\{\xi\in A:\xi(0)=\xi_0\})=\theta P^\xi_{\xi_0}(A)$ where $P^\xi_{\xi_0}$ is the law of the time reverse of the random walk $(\xi(s))_{s\geq 0}$ with transition rates $a(\xi,\xi')$ starting at ξ_0 . We then note that $P^\xi_{\xi_0}$-a.s. $\bar x_{\xi(s)}$ is a stationary process (with respect to the equilibrium law for $\bar x$). Then (4.59) follows immediately from the ergodic theorem. This completes the proof.

Acknowledgement We thank P. Pfaffelhuber for many comments and suggestions while the manuscript was in preparation and Zhenghu Li for suggestions for the proof of Lemma 1.8. We thank the referee for very careful reading of the manuscript.

References

[BCGH] J.B. Baillon, Ph. Clément, A. Greven, F. den Hollander: On the attracting orbit of a nonlinear transformation arising from normalization of hierarchically interacting diffusion Part 2: The non compact case, Journal of Functional Analysis 1997.

[CFG] T. Cox, K. Fleischmann, A. Greven: Comparison of interacting diffusions and an application to their ergodic theory. Probability Theory rel. Fields **105** (1996) 4, 513-328.

[CG1] J. T. Cox, A. Greven: Ergodic theorems for systems of locally interacting diffusions. Ann. of Prob. **22**(2). 833-853 (1994).

[CG2] J. T. Cox, A. Greven: The finite systems scheme: An abstract theorem and a new example, Proceedings of CRM Conference on Measure valued processes. Stochastic partial differential equation and interacting Systems. CRM and Lecture Notes series of the American Mathematical Society. Vol. **5**, 55-68 (1994).

[CGS1] J.T. Cox, A. Greven, T. Shiga: Finite and infinite systems of interacting diffusions. Prob. Theory and Rel. Fields, **103**, 165-197, 1995.

[CGS2] J.T. Cox, A. Greven, T. Shiga: Finite and infinite systems of interacting diffusions. Part II. Math. Nachrichten **192** (1998).

[D] D. A. Dawson: Measure-valued Markov Processes, In Ecole d'Eté de Probabilités de Saint Flour XXI, Lecture Notes in Mathematics **1541** 1-261, Springer, 1993.

[DG0] D. A. Dawson, A. Greven: Hierarchical models of interacting diffusions: Multiple time scales, Phase transitions and cluster-formation. Prob. Theor. and Rel. Fields, **96**, 435-473, 1993. (1993).

[DG1] D. A. Dawson, A. Greven: Multiple space-time scale analysis for interacting branching models. Electronic Journal of Probability, Vol. 1, Paper no. 14, 1-84, (1996).

[DG2] D. A. Dawson, A. Greven: Hierarchically interacting Fleming-Viot processes with selection and mutation: Multiple space time scale analysis and quasi-equilibria, Electronic Journal of Probability, Vol. 4, Paper 4, 1-81, (1999).

[DGV] D. A. Dawson, A. Greven, J. Vaillancourt: Equilibria and Quasi-equilibria for infinite systems of Fleming-Viot processes. Transactions of the American Math. Society, Vol. 347, Number 7, (July 1995).

[DH] D. A. Dawson, K. Hochberg: Wandering random measures in the Fleming-Viot model, Ann. Probab. 10(1982), 554-580.

[DP] D. A. Dawson, E. A. Perkins: *Historical Processes*, Memoirs of the American Mathematical Society, vol. 93, no. 454, (1991)

[DP2] D.A. Dawson and E.A. Perkins: Long-time behavior and coexistence in a mutually catalytic branching model, Ann. Probab. 26 (1998), 1088-1138.

[EK] S. N. Ethier, T. G. Kurtz: Markov Process - Characterization and Convergence, John Wiley and Sons, (1986).

[EP] S.N. Evans and E.A. Perkins: Measure-valued branching diffusions with singular interactions, Can. J. Math. 46, 120-168.

[FG] K. Fleischmann, A. Greven: Time-space analysis of the cluster-formation in interacting diffusions. Electronic Journal of Probability, Vol. 1, Paper no. 6, Seite 1-46 (1996).

[G] A. Greven: Phase Transition for the Coupled Branching Process, Part I: The ergodic theory in the range of finite second moments. Prob. Theory and Rel. Fields **87**, 417-458 (1991).

[KS] I. Karatzas and S.E. Shreve: Brownian Motion and Stochastic Calculus, 2nd ed., Springer-Verlag (1991).

[K] U. Krengel: *Ergodic Theorems.* de Gruyter Studies in Mathematics 6, Berlin-New York, (1985).

[LS] T. M. Liggett and F. L. Spitzer: Ergodic theorems for coupled random walks and other systems with locally interacting components, Z. Wahrsch. verw. Gebiete, **56(4)**, 443–468 (1981).

[P] E.A. Perkins: Measure-valued branching diffusions with spatial interactions, Probab. Theory Relat. Fields, 94 (1992), 189-245.

[P2] E.A. Perkins: Conditional Dawson-Watanabe processes and Fleming-Viot processes, Seminar in Stochastic Processes, 1991, Birkhauser, 142-155.

[R] P. Révész, *Random Walk in Random and Non-random environments*, World Scientific, (1990).

[RW] L.C.G. Rogers and D. Williams: *Diffusions, Markov Processes and Martingales*, volume 2, John Wiley and Sons, (1987).

[RY] D. Revuz and M. Yor: *Continuous Martingales and Brownian Motion*, Springer-Verlag, (1991).

[S] T. Shiga: Ergodic theorems and exponential decay of sample paths for certain diffusions, Osaka J. Math. **29**, 789-807, (1992).

[SS] T. Shiga and A. Shimizu: Infinite-dimensional stochastic differential equations and their applications. J. Math. Kyoto Univ., **20**, 395-416 (1980).